DOCUMENT RESUME

ED 102 162 95 SP 008 948

AUTHOR	Houston, W. Robert, Ed.
TITLE	Competency Assessment, Research, and Evaluation.
INSTITUTION	Multi-State Consortium on Performance-Based Teacher Education, Albany, N.Y.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.; Office of Education (DHEW), Washington, D.C.
PUB DATE	Mar 74
NOTE	273p.; Report of a National Conference (University of Houston, Houston, Texas, March 12-15, 1974)
AVAILABLE FROM	American Association of Colleges for Teacher Education, Suite 610, 1 Dupont Circle, Washington, D.C. 20036 ($3.00); National Dissemination Center for Performance Based Education, 150 Marshall St., Syracuse University, Syracuse, New York 13210 ($3.00)
EDRS PRICE	MF-$0.76 HC-$13.32 PLUS POSTAGE
DESCRIPTORS	*Conceptual Schemes; *Educational Research; Organizational Change; *Performance Based Teacher Education; Professional Training; *Program Effectiveness; Social Factors; Teaching Techniques

ABSTRACT

This book is a compilation of papers that consider the current state of the art of competency/performance-based teacher education. Organized into six parts, the first part of this book reports three frames for conceptualizing research in competency-based education. The keynote address outlines five major areas for research: assessment of professional competency, program effectiveness, effectiveness of instructional strategies, social context, and institutional change. Parts two through five focus on these areas of research. (Editor/PD)

COMPETENCY ASSESSMENT, RESEARCH, AND EVALUATION

A Report of a National Conference
March 12-15, 1974

BEST COPY AVAILABLE

This publication is being distributed by the Multi-State Consortium on Performance-Based Teacher Education as a function of one of its major objectives -- "the dissemination and communication of information about performance-based teacher education."

The opinions expressed herein should not be construed as representing the opinions of the United States Government, the Multi-State Consortium on Performance-Based Teacher Education or any of the states which are members of the Consortium, or the National Dissemination Center for Performance-Based Education, Syracuse University.

BEST COPY AVAILABLE

COMPETENCY ASSESSMENT, RESEARCH, AND EVALUATION

A Report of a
Stanford Conference
at
University of Houston
March 13-15, 1974

Sponsored by:
The Multi-State Consortium on Performance-Based Teacher Education
The National Commission on Performance-Based Education
The National Consortium of CBE Centers
The National Institute of Education, Program on Teaching and Curriculum
The Committee on Performance-Based Teacher Education,
Association of Colleges for Teacher Education
Division of Education Systems Development, United States Office of Education
Professional Development Center, University of Houston
Teacher Corps

Copies of this report may be obtained for $3.00 from:

American Association of Colleges
for Teacher Education
Suite 610
One Dupont Circle
Washington, D. C. 20036

National Clearinghouse Center for
Performance-Based Education
130 Washington Street
Syracuse University
Syracuse
New York 13210

This Conference was sponsored and this report was published as part of grants from the U. S. Office of Education and National Institute of Education, Department of Health, Education, and Welfare. The opinions and other content do not necessarily reflect the position or policy of the Agency, and no official endorsement should be inferred.

CONTENTS

COMPETENCY ASSESSMENT, RESEARCH, AND EVALUATION

FOREWORD — R. Robert Houston, Conference Director and General Editor of this Report — Page[illegible]

PART I — [illegible]

Chapter 1 — [illegible] — [illegible]

Chapter 2 — [illegible] — [illegible]

Chapter 3 — [illegible] — [illegible]

PART II — [illegible]

Chapter 4 — [illegible] — [illegible]

Chapter 5 — [illegible] — [illegible]

Chapter 6 — [illegible] — [illegible]

Chapter 7 — [illegible] — [illegible]

Chapter 8 — [illegible] — [illegible]

Chapter 9 — [illegible] — [illegible]

FOREWORD

PART I

FRAMEWORK FOR RESEARCH AND DEVELOPMENT

1 [illegible]

[illegible]

BEST COPY AVAILABLE

based education ought to be the style and strategy for training teachers is part logical and part psychological. The logical aspect relates to the structure, amount, and kinds of evidence needed to constitute sufficient proof that the theory or model or strategy is useful. The psychological aspect relates to what individuals will demand because of their intuitive and preferential reactions to the concept. The psychological aspects of belief and acceptance are not totally responsive to the logic of data.

One component of the psychological aspects of stimulating conviction is that as competency based teacher education catches on, significant social implications for the teaching profession will be apparent. As a consequence, people whose lives will be affected by competency based education are looking at the political implications of certifying teachers on the basis of demonstrated competence. Their concern is twofold: first, that the concepts of competence will be defined independently of the conditions of teaching; second, that the criteria of competence will not take into account other factors that influence both teaching and learning.

The conclusion to be drawn from this concern is that research designs must include the variety of factors that probably impinge on teaching performances and their consequences in learning. It is reasonable to assume, for example, that variations in administrative structure, curriculum, and characteristics of children and teachers affect teaching performance itself and also learning. It is obvious that an important research problem in its own right is to measure the relative effects of such factors. Descriptions of competence divorced from the conditions of teaching and learning are not likely to be believable.

The criteria for assessing the sufficiency of evidence appear to be of two kinds. One set is related to the validity of the results of the research—what data support a hypothesis that a teaching performance is effective, and how generalizable are the results. The second set are more elusive, related more to the context and content of the research. The research must be believable in that it attends to the complexities and realities of teaching.

I address myself primarily to the kinds of knowledge needed to support the proposition that competency training is the most effective way to educate teachers. I will discuss several basic research problems whose solution is required if the effectiveness of competency training is to be demonstrated and accepted. Other speakers will discuss these problems in greater detail, elaborating their conceptual and methodological difficulties.

Let me begin with what I think is the basic problem, which happens also to be fraught with implications of the kind previously described, that is, what is the nature of teaching competence? Some of us seem to assume that because we have more precise descriptions of teaching competencies that we have resolved the question of what is teaching competence. One of the most commonly stated accusations aimed at competency education is that we believe that a list of behavioral descriptions of competence constitutes a description of competence. We are quite properly asked, "How do you know that if the teacher performs as described that the performance will have a desirable effect on student learning?"

A competence may be described as a behavior of a teacher; for example, "asks students to give reasons for their opinions on controversial topics

BEST COPY AVAILABLE

being discussed in class." Presumably, a high frequency of such statements indicates that the teacher has the behavioral performance in his or her repertoire. But is such information evidence of competence? It is only if we assume that the appearance of the behavioral act is synonymous with competence.

Such an approach is rightly charged, in my judgment, with indifference to the consequences of the event. The more cynical of critics point out that this concept of competence leads logically to certifying as competent teachers who emit the behavioral performances in the absence of students. While such criticisms are unnecessarily harsh, they point to the problem—is competence to be defined in terms of effects manifested in students, in terms of observable acts of teachers, or in terms of pairs of acts and their effects?

It seems to me that we can no longer avoid saying that teaching competence must be defined in terms of effects on pupil learning. Unless the research and development movement behind competency based education addresses the problems implicit in that concept of competence, it seems to me there is no hope of ever persuading anybody that the competency based education movement is doing anything other than developing semantic descriptions of what constitutes competence. I propose that the first and most important research problem to attack is the question of what teaching performances produce significant effects on learning.

There are two general research methodologies that may be used for this purpose. I would like to describe them very briefly. The first of these is an inductive approach to the analysis of the relationship between teaching performance and student learning. This methodology proceeds as follows. The research should be conducted so that actual teaching may be observed, preferably that of more experienced teachers. You need not have preconceptions about what kinds of performances are likely to affect student learning, or your observations of performance may be guided by theory or previous research. But you must have a clear and specific conception of the kinds of student learning that you are going to study.

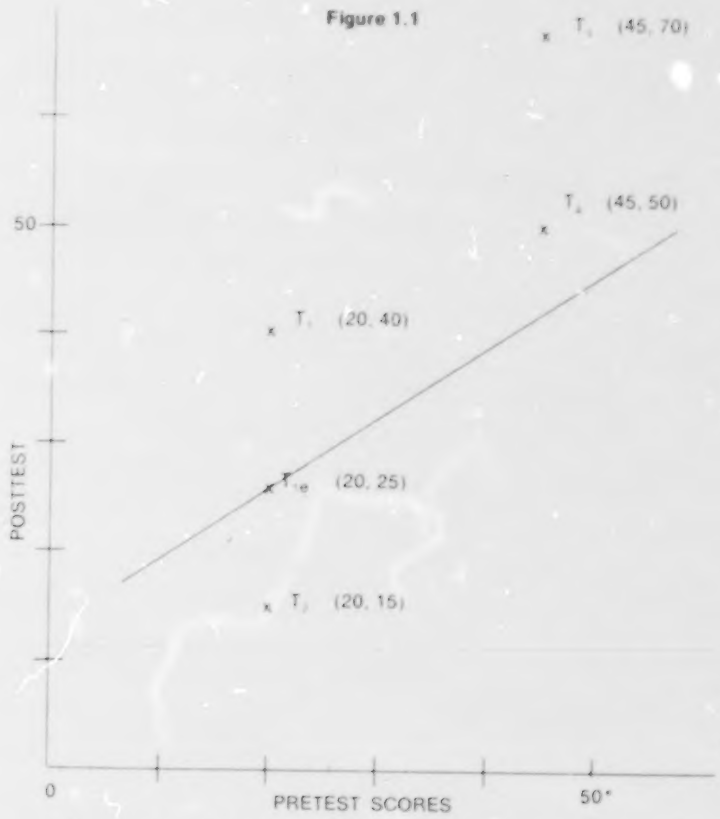
The simple form of the methodology consists of measuring student learning at two points in time.* Given two such measurements obtained on each of the students in the classes of a sample of teachers, a regression of the later or posttest scores is performed on the earlier or pretest scores. The results of this statistical analysis can be diagramed as in Figure 1.1.

The points are represented by the mean scores for a teacher's class, the first number being the mean for the pretest, the second, the mean for the posttest. The line through the points is called a regression line and points on it represent expected scores. Given the information in the analysis, we expect T_1's posttest mean score to fall on the line at T_{1e}; similarly, we would expect T_2's posttest mean to be located on the line at the same point. Both posttest means, however, locate T_1 above the line and T_2 below the line.

*Learning in this speech is an omnibus category meant to include all desirable changes in students, cognitive, affective, and social. Nor am I concerned about the method of measurement or the difficulties of measuring some changes. The researcher needs to be specific about the changes desired and have developed reliable and valid ways of measuring them.

BEST COPY AVAILABLE

Figure 1.1



Regression Analysis of Posttest Scores on Pretest Scores Using the Scores of the Students of a Sample of Teachers

*The metric is arbitrary. The first number in parentheses represents the mean score for the T's class on the pretest, the second number, the mean on the posttest.

When you do this type of analysis, you are looking for teachers who depart markedly from the regression line because you want to study them intensively to find how they differ from teachers who are closer to the regression line. The general idea is that teachers whose student performance shows marked departures from the line may be highly effective, other things being equal. We ignore here the many technicalities of this type of analysis to make the point that the method provides a way of identifying teachers who may be unusual and provides a way of inducting hypotheses about what constitutes performance related to student learning.

BEST COPY AVAILABLE

The next step is to study the observed performances of these teachers and find the performances which discriminate their teaching performances from that of other teachers in the analysis. If you gather enough information about the performance of the children and the performance of the teacher, such as the kinds of schools in which they are, the composition of the class, and similar factors, you stand a fairly good chance of sorting out those performances associated with the act of teaching which distinguish more effective teachers from other teachers.

It is not true, however, that these performances are guaranteed to be the ones that are associated with effective teaching. Let us consider the problems of this type of a study with respect to the two questions that I asked earlier: what constitutes evidence, and how much evidence is sufficient? At the end of this type of study, all you have is a set of hypotheses about the relationship between teaching performance and student learning that may have some validity. These hypotheses must be tested either by doing experimental studies or by doing replications of the original study or both.

The problem of how much evidence is sufficient arises from the complexity of these studies. It is feasible to study a limited number of student outcomes. To gather data on students on relatively few outcomes requires several days of testing. To study the teachers over time requires multiple observations. The question is, how many studies of this kind are needed before we have sufficient evidence of the components of competence? Where are they to be located in a matrix of pupil changes, types of schools, differences in organizational structures, differences in levels of education, and differences in teacher and student characteristics before we have confidence that we have working hypotheses that describe competence?

What we must not do is propose or claim that we need 500 of these studies and that it will take 10 years and $10 billion of research money to describe teaching competence. We must recognize that we are living in an environment of pessimism about the probability of demonstrating that there is a relationship between teaching competence and pupil learning. One of the goals of the research and development program on competency based training should be to demonstrate the probability that teaching competence can be identified. It seems to me that several studies of the kind that I have described are likely to demonstrate some relationships. Such studies will at least eliminate some hypotheses about teaching or indicate weak relationships. Because such studies involve correlational designs, however, and because they are inductive in character, potentially valid hypotheses may not be discovered. Another type of strategy is needed that is theoretical-experimental in character.

A theoretical-experimental type of study varies teaching performances systematically and assesses the effects of these variations on student performance. The manipulation of the performances may be of two kinds. One manipulation is to vary the degree of the skill of a performance to determine the relation of the degree to the amount of the effect. Such experiments provide the data necessary to establish a criterion level for the performance.

A second type of experimental manipulation requires varying the kinds of performances to assess their effects on the same outcomes. Such studies

BEST COPY AVAILABLE

provide comparative data on the effects of performances, their interchangeability, and their interdependence.

These experiments should be conducted under controlled conditions where the conditions and materials for teaching are identical and where students are assigned randomly to classes. They may also be conducted under field conditions using fractional factorial designs. This type of design involves selecting relevant contrasting conditions in which the teaching occurs such as those associated with the characteristics of students, schools, or curricula. Fractional factorial designs are notably efficient because relatively few teachers can be used to test a relatively large number of hypotheses. When used under field conditions, they provide tests of interactions among school and pupil factors and teaching performances.

The advantage of experimental designs is that they provide a way of testing hypotheses systematically. Experiments are powerful ways of testing theories of teaching by testing the hypotheses derived from them. One can begin as follows: select one or more models of teaching, define or describe the kinds of pupil changes that the theory predicts will occur if children were instructed by teachers applying that model, and derive hypotheses about what kinds of teaching competencies produce these changes in students. The experiments are then designed to test these hypotheses.

The advantage of using theories or models as starting points for deriving hypotheses is that the size of the hypothesis pool will be greatly enlarged. The disadvantage is many hypotheses will be derived of unknown validity either because they are loosely derived from the theory and have no empirical support from any source or because the theories themselves have weak empirical support. At this stage of research on teaching, it is probably better to include more weak than less hypotheses in the pool for testing.

A research and development strategy on teaching competence ought to include both inductive correlational and theoretical-experimental studies. Further, each inductive correlational study ought to be followed by experiments which test the validity of the hypotheses derived inductively.

The problem remains, what constitutes sufficient evidence to make reasonably definitive statements about teaching competence? One approach to answering this question would be to consider the consequences of training for a set of competencies that has been derived from an inductive correlational study. Here the evidence is relatively weak—no replication of the original study and no experimental verification. There seems to be no risk in using such information, particularly if the training program studies what happens when its teachers use the competencies in real teaching situations. In this case, the opportunity exists to generate more data about the effects of the teaching performance.

The problem of sufficient evidence seems to lie in the political arena. If a state agency requires that programs train for the competencies derived from even a few inductive correlational studies which have not been replicated, one has cause for concern.

It also seems unwise to assume that there will be an end-point at which it can be said that we now have sufficient evidence to "justify" competency training. Rather the hypothesis-generating, hypothesis-testing life style should characterize the development of the competency based movement.

BEST COPY AVAILABLE

There ought not to be a limit about all most important research on teaching competence.

A few general comments about the problem of studying teaching competence are relevant here. In studying teaching competence we look to associate out systemic effects on students from teacher effects. Many of the goals of education are broad. Changes will invariably in most occur as the consequence of the teaching of the different kinds of teachers which the pupil will confront and also as a consequence of other kinds of influences such as those from peers. It seems to me that our main problem in generating a rigorous approach should be to find a way to generate it in such a way that the systemic effects are treated as contextual moderators of teaching performances. Research on competence ought not to be, at least initially, research on effects that occur as the consequence of that piece of education. Such research is both interesting and important but it is not really related to demonstrating the effects of teaching competence. When we know more about the specific effects of teaching acts under known conditions, we will be prepared to study systemic effects.

It is important to keep in mind that in conducting the type of controlled experiment research that I referred to earlier. It is possible to design those studies so that a wide variety of system effects may be studied as interacting with the effects of teaching performances. You can determine the kinds of influences that the administrative climate of the school, for example, has on teaching performance and how performance varies within different organizational structures or within different types of programs. But our problem is to account for, or to describe how much of the variance in student performances actually can be accounted for by various teaching performances. The goal of our research is to find the effects of teaching. other factors in this research are important if they influence or so affect learning that the influence of teaching is minimal. If we cannot determine the effects of teaching performances, the competency based movement is irrelevant.

Another problem is that we need to know how different kinds of teaching performances cluster - we know practically nothing at all about the structure of teaching performances. There is some evidence that specific descriptions of performance cluster in factors. Most descriptions of competence are lists of behavioral acts. Interdependencies among them are made highly. But if such specific acts tend to occur together in groups, and if these clusters have demonstrable relations to changes in students, then these clusters are probably the basic units of teaching performances.

Then the question is, what is competence? Is it the performance enhanced in such factors, or is it the basic descriptors that are applied not specifically and that an observer may check as present or absent in a teaching episode? We really do not know, but some evidence suggests that perhaps it is these clusters and not the most minimize aspects of the total performance that produce learning effects. The inductive correlational type of study provides the empirical data required for the analysis needed to study this problem.

A fundamental problem facing us from a research point of view is: what is the unit of competence? Is it the simple semantic descriptor which is

2. National Institute of Education and Research in [illegible]

BEST COPY AVAILABLE

catalog of funded projects is usually available, it is rarely possible to obtain a catalog of proposed projects.

Establishment of an Alternative Relationship

It is possible to create a relationship between funding agencies and their related research communities based on the foregoing four kinds of initiatives. An approach is being tested in the Program on Teaching and Curriculum (directed by Garry L. McDaniels) and in the Essential Skills Program (directed by Marshall Smith) at the National Institute of Education. After the general approach is described, the specific initiatives being taken will be presented. The general approach has four elements: (a) specifying major problems, (b) obtaining simultaneous input from field researchers regarding possible projects, (c) establishing alternative strategies for attaining goals, and (d) communicating alternative strategies with project recommendations.

Specifying Major Problems

Instead of having the funding agency identify major problems at a broad or abstract level, the agency must force concrete definition of problems. A Federal funding agency cannot remain at the general level of the goals of the authorizing legislation, just as the private foundation cannot effectively use the broad goals identified in its request to the Internal Revenue Service for tax-exempt status. Also, if the character of the problems are not fully described, it is also rather difficult to assess the accomplishments of the effort.

Obtaining Simultaneous Input from Field Researchers

Investigators are storehouses of project ideas. As time and conditions allow, an idea is selected, refined, and sent to a funding source by an investigator. Three problems result:

- The ideas are sent from different people at different times
- It is difficult to relate projects across agencies and foundations
- Each person's proposal is developed in a time lag from the recent work of other investigators; therefore, he may plow the same ground.

If cross-project analysis is to be possible, it is essential that simultaneous input from field researchers be achieved.

Establishing Alternative Strategies

Not only must the input from field researchers be obtained simultaneously, it must also be obtained in a manner which identifies the intellectual and administrative relationships among projects. The development of a plan requires that the potential research ideas be organized into a logical agenda. Further, the organization should demonstrate how the projects relate to the achievement of an objective worthy of public or other support.

Communicating with the Research and Development Public

Finally, the steps described above provide information to be communi-

BEST COPY AVAILABLE

cated to the research and development public. Much of that public will be informed as a result of having been a part of the process. The remainder will have the time required for receiving and analyzing the pertinent documents.

The National Conference on Studies in Teaching

To further clarify the implications of these elements, the steps being taken in the Program on Teaching and Curriculum will be described. In regard to specification of major problems, the staff is meeting with students, teachers, administrators, teachers in colleges of education, etc. to get a summary of problem statements. These problem statements will serve as a charge to the researchers who will make project recommendations during the planning conference. Perhaps it would be ideal to have everyone meet simultaneously. Such a meeting would, however, require a quota system for representation of the relevant groups—which would result in huge working sessions that would be impossible to manage.

To obtain simultaneous input from field researchers, the present conference has been planned. To establish the organizational framework for the development of the plan, the panel chairmen will spend several weeks in making a separate analysis of subject area objectives, in cooperation with the conference participants. During the conference, the participants will refine those approaches, identify possible areas of research to implement the approaches, and specify the research projects which should be considered by the Agency.

Finally, to communicate these proceedings, several steps will be taken. First, all participants will leave the conference with complete volumes of panel suggestions. These will be edited for minor errors and sent to the Government Printing Office for printing and distribution. The analyses by the NIE which result in strategy recommendations will be available for general distribution at the end of August 1974. The logistics of printing and distribution have been established.

Substantive Form of the Panels

Preliminary discussion by NIE staff and consultants since August 1973 has fielded a formulation of the substantive structure of the conference—and the consequent foci of its 10 panels. In part, this structure is based on a categorization of major concerns in the career of the teacher: (a) recruitment, selection, and retention; (b) training, performance, and effectiveness; and (c) utilization. In addition, three foremost concerns of all Research and Development in teaching will be taken into account by panels on Personnel Roles in New Instructional Systems, Research Methodology, and Theory Development. Thus, the Conference will consist of the following panels:

1. Panel on Teacher Recruitment, Selection, and Retention

2-6. Panels on Teacher Training, Performance, and Effectiveness

   2 Panel on Teaching as Human Interaction

   3. Panel on Behavior Analysis in Education

   4. Panel on Teaching as Skill Performance

BEST COPY AVAILABLE

5. Panel on Teaching as a Logical and Linguistic Process
6. Panel on Teaching as Clinical Information Processing
7. Panel on Teacher Utilization
8. Panel on Personnel Roles in New Educational Systems
9. Panel on Research Methodology

In the following paragraphs, each of these panels is briefly characterized.

Panel 1: Teacher Recruitment, Selection and Retention

The panel on Teacher Recruitment, Selection and Retention will be concerned with planning Research and Development on devising, validating, and installing methods for improving formal and informal decisions at various points in a teacher's career:

(a) admission to a teacher education program
(b) retention in the program
(c) admission to student teaching or internship
(d) graduation from the program
(e) certification or licensure
(f) employment by a school system
(g) granting tenure
(h) promotion in rank and pay

The classical paradigm in Research and Development on recruitment, selection, and retention consists of identifying criteria, selecting potential predictors of those criteria, and determining empirically their predictive validity. The panel will accordingly pay much attention to the matters of criteria, predictors, and ways of basing decisions on their relationships. The criterion problem will entail consideration of competency assessment and teacher evaluation procedures.

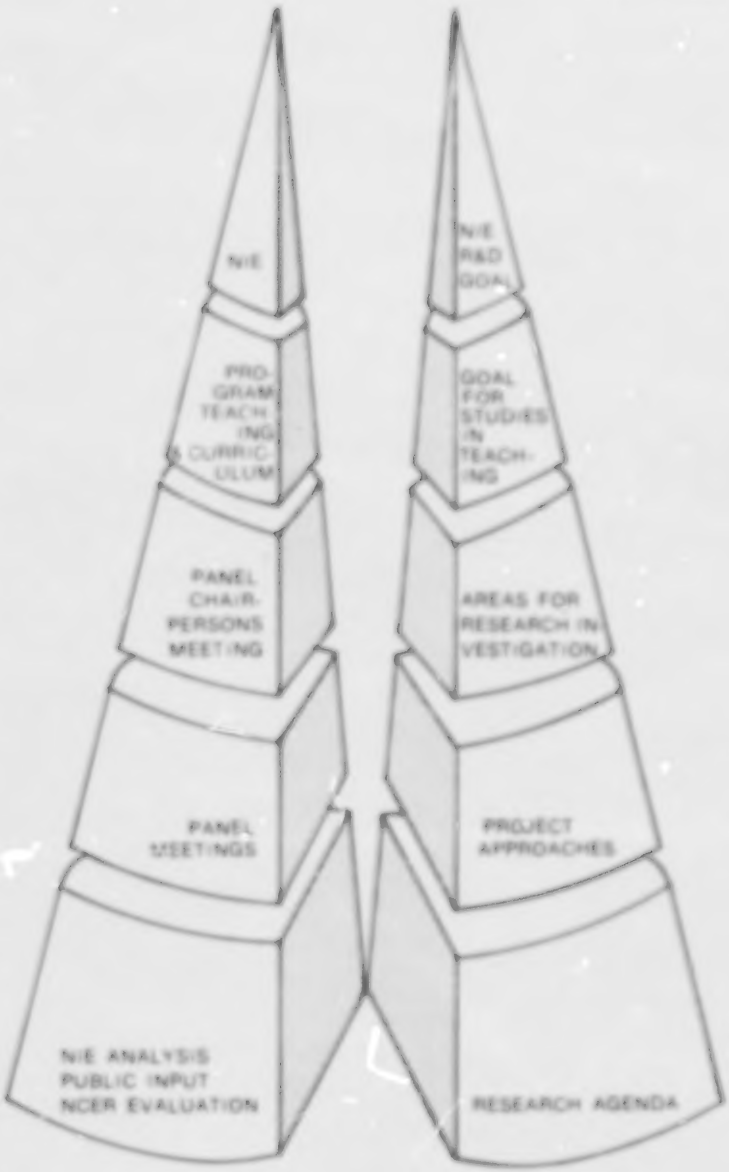
The panel will also deal with self-selection—ways in which persons are attracted toward or turned away from teaching, i.e., the occupational choice and attrition processes. It will examine the social class, ethnic, and other origins of teachers and the differences these make. It will also deal with the possibility that success in a training program is more a function of the program than of the candidate and his characteristics.

Panel 2: Teaching as Human Interaction

Many studies have dealt with observed human interaction (either live or recorded) as independent and dependent variables. In some cases, observations of human interaction have been made merely to describe differences between kinds of teachers in various settings and contexts. These studies have been reviewed in several monographs. This approach has been the most frequently used in research on teaching during the past two decades.

To review the current state of this approach, to lay out the most promising next steps (such as projects on taxonomies of teacher behavior, chains and patterns of teacher and student moves, computerized storage and analysis, and archival banks of videotaped interaction records), and to formulate

BEST COPY AVAILABLE



BEST COPY AVAILABLE

those projects in optimally definite terms -- these illustrate the tasks that can be undertaken by a panel in this highly cultivated, but far from exhausted, field.

* Research and development on teacher training, performance, and effectiveness deals with (1)either variables (teacher characteristics and behaviors) in either or both of two ways: (A) as dependent variables (with teacher training methods serving as independent variables), and

(B) as independent variables (with effects on students serving as dependent variables).

Work in this area has been based on a number of paradigms, not a relatively independent, hence several corresponding panels (2-6) will spin here.

Training is construed to embrace acculturation processes that infuse the trainee with the values and norms of the profession. It includes both preservice and inservice training. These panels will emphasize aspects of training that require practice, i.e., training intended to provide "knowledge how" not merely "knowledge that." The panels will be encouraged to develop new paradigms, cutting across or departing radically from present ones.

Panel 3 Behavior Analysis in Education

Somewhat more recently, but with equal vigor and productivity, the behavior analysis in education approach has been applied to the improvement of teacher training performance, and effectiveness.

The success of teacher training and especially of classroom management by these methods has been frequently reported. Thus, this thriving line of attack on several important problems in teaching and teacher training merits the attention of a panel. This panel might seek to increase the attention already being given to private events and cognitive processes. It could advance the effort to apply behavior analysis in education to the achievement of humanistic educational objectives. It could consider studying the effect of students on teachers through the cues provided by students, perceived by teachers, and responded to by teachers, and thus the ways in which students may reinforce teachers' responses.

Questions have been raised about the morality of "controlling" behavior through behavior modification methods. Hence, this panel may undertake to address such ethical issues.

Panel 4 Teaching as Skill Performance

During the past decade, an analytic approach to the training and evaluation of teachers has frequently been adopted. In such an approach, teaching is broken down into component skills, such as questioning, explaining, listening, and reinforcing. Little consensus has been reached on what skills should be identified. But the approach itself has flourished, lying at the basis of such innovations as microteaching and minicourses, in which teachers are trained in only one or a few skills at a time.

The analytic approach has also led to the movement to develop teacher-training products. Such products call upon the trainee to practice a given skill, in addition to reading, listening, or watching material about the skill. Hundreds of such products have been developed and cataloged, in part as a response to the nationwide attempt to make teacher training and evaluation "performance based." Their promise will be realized, however, only to the degree that they benefit from further research and development. The analyses should become more valid, logically and psychologically.

BEST COPY AVAILABLE

Panel 5 Teaching as a Logical and Linguistic Process

Studies of linguistic and logical aspects of classroom interaction here and in England have varied widely in the particular aspects of language studied, the kinds of classrooms selected for research, and the methods used for data collection and analysis.

Aspects of language studied have included the logical operations engaged in by teachers and pupils, the types of pedagogical moves in the classroom interaction "game" and their distribution among teachers and pupils, formal speech registers used in the school, as contrasted with the home, and their various special characteristics in various school subjects, and the sociolinguistic rules operating in the classroom seen as a special setting for face-to-face interaction.

Most research has been conducted in teacher-structured recitations, less in small-group and pupil-led discussions and group projects. Many researchers have assumed a set of norms governing classroom interaction that is shared by all participants; others have deliberately worked in multiethnic settings and considered variations in these norms as a special problem for analysis.

On the basis of this previous research activity, this panel will undertake to formulate the direction in which research should move to provide improved understanding of logical and linguistic phenomena in classroom settings and the ways in which that understanding might be used to improve the teacher's work.

Panel 6 Teaching as Clinical Information Processing

This panel should be concerned with the thought processes that precede the teacher's choices of action. How teachers think, make judgments, size up situations, plan and react to unanticipated contingencies—all these should be of concern here. Thus, this panel would be concerned with the study of teaching as thinking, problem solving, judgment, and decision making. It might also consider the value of a clinical (possibly anthropological) approach to research on teaching. Studies of the teacher as clinician would be concerned with the way in which theory, although it does not prescribe for practice, nonetheless informs practice.

The diagnostic approach to teaching and teacher preparation would also be concerned with the teacher's expectations. Such expectations have been found to influence not the student's intelligence, as was once claimed, but teacher behavior, teacher-student interaction, student motivation, and student achievement of the objectives at which instruction is aimed. This panel should, therefore, give attention to the determinants and effects of teachers' expectations concerning students in various categories (sexes, races, ethnic groups, age-levels, etc.).

Panel 7 Teacher Utilization

After teachers have been selected and trained, they go to work in a school which is part of a school system. What happens to them in the months and years of their work in the school obviously influences greatly the ways in

which they learn, how and what is most important, influence their students.

The term *university* refers to the ways in which teachers are required or permitted to work by the administrative, organizational, and physical circumstances that their school and school system provide. The many new ways of defining teachers in this sense include team teaching, teacher participation, nongraded classrooms, school without walls between classes, accountability, new merit-based teacher evaluation systems, voucher plans, teacher unions, flexible scheduling, and differentiated staffing. How is academic talent located? What are the causes and effects of what is often considered to be a new level of competency in mathematics and secondary science? Should curriculum variation be used in studies of school outcomes so that their value can be evaluated? Can instructional hardware systems be arranged so as to improve teaching and motivation teacher education?

Sociologists, school administrators and teachers' organizations have been especially active in developing and analyzing innovations in teacher preparation. The panel here should contain representatives of these groups.

Panel II: Personnel Roles in New Instructional Systems

Some writers regard the traditional classroom organization of the school and the teacher's role in that classroom as structural. During the next decade or two, they presently prevalent arrangements will be supplanted by what are now regarded as revolutionary new arrangements. The new schools will incorporate instructional systems that fully exploit the potentialities of computers, television, and various other products of recent technology.

The role of the teacher will change in correspondingly radical ways. In some views, the teacher will have much less to do with the immediate affairs of facilitating achievement of cognitive objectives because computer assisted instruction will carry much more of that burden. Instruction will be more effectively adapted to the needs and capabilities of each individual student.

Such individualization will absorb the labor, the uncertainty, the initial phases, the rehearsal efforts, and the trials-and-errors of instruction so as to optimize learning of the individual student. Beyond these kinds of attainment, the new systems will assign the subject-matter preparations to the individual's aptitudes, learning styles, activities, preferred types of reinforcement, and immediate past history of responses. They will also use new approaches to the evaluation and grading of student achievement, turning more to criterion-referenced approaches to cure. These approaches will lead to the objective of greatly increasing, indeed maximizing, the percentage of students that master all aspects of all their courses. Such mastery levels will be achieved much more abundantly in the new instructional systems.

What will be the teacher's role in the new systems of instruction? Will the teacher become less the curriculum planner, the diagnostician and prescriber, the evaluator? Will the teacher become more group process facilitator, the mediator of emotional responses, the model for moral and esthetic development?

This panel will develop approaches to the research and development that will provide bases for predicting and optimizing the roles of various kinds of

3 Closing the Knowledge Gap

[illegible]

BEST COPY AVAILABLE

opposed to a form that relies upon syllabi, readings, discussions, and the like; and the costs and benefits of including in a program elements that deal with understanding self as teacher, interpersonal awareness and sensitivity, and skills in group problem solving. Without such information, decisions as to program design and operation must also be based on something other than firm knowledge.

As in the case of policy oriented research, two conditions must exist before practice oriented research can be carried out. First, programs must be in operation that reflect the practice of concern. Second, either programs must be found that are comparable in structure, organization, and student body and are willing to compare alternative practices, or a program must be found that is willing to design and carry out alternative practices under experimental conditions while continuing to carry out program operations.

In some respects, these conditions resemble those needed to carry out policy oriented studies. The existence of fully operational programs is a case in point. In other respects, however, practice oriented studies go beyond policy oriented studies in their demand upon control over unwanted sources of variation. In this respect they become more difficult to carry out, for they impose considerably greater constraints upon programs that are taking part in the research.

Given these conditions, practice oriented studies in the arena of CBTE will probably take the form of "systematically manipulated" field experiments, in contrast to the "natural" field experiments that characterize policy oriented studies.

By viewing the requirements for practice oriented studies in this light, both policy makers and practitioners should be able to understand why research of this kind has not been undertaken thus far in the history of the CBTE movement. Simply put, it is for the same reason that policy studies have not been undertaken: there have not been programs sufficiently well developed to permit such research to be carried out. Now that CBTE programs are beginning to reach maturity, research of this kind becomes possible.

Intra-Institutional Studies

Since the effects of program practices and procedures would be expected to interact powerfully with learner characteristics and the kind of learning outcomes desired, it is likely that most practice oriented research will be carried out within the context of a particular CBTE program. This will permit as many sources of unwanted variation in such studies to be controlled as possible. The conditions outlined earlier, however, must be obtained before a CBTE program would qualify as a site for such research. Paraphrased, these conditions would include a fully operational CBTE program where personnel of the program are interested in and able to carry out experiments on the effects of systematically designed program variations while keeping constant the learning outcomes desired from the program. In addition, a participating institution must have a reasonably large number of students enrolled in its CBTE program; it must be organized in such a way that bona fide experimental and control groups can be established; it must be organized in such a way that true component variations are able to be im-

Table 3.2

OUTPUT MEASURES, SOURCES OF VARIATION TO BE CONTROLLED, AND COST ANALYSES TO BE PERFORMED IN PRACTICE ORIENTED STUDIES

Output Measures		Sources Of Variation to be Controlled	Cost Analyses to be Performed
Short Term	Long Term		
Whatever measures of teaching competency a particular program has adopted as the criterion for exit from either the program as a whole or segments within it (these may be developed as a by-product of participation in particular research studies) Other measures of impact on program operation, e.g., acceptability, organization and time requirements, resource requirements	None	Initial differences in treatment groups Unwanted or unexplained treatment conditions Lack of equivalent contexts for competency demonstration	Instructional costs Competency assessment costs Program management costs Program governance costs

BEST COPY AVAILABLE

BEST COPY AVAILABLE

plemented; and high quality measures of learning outcomes must be obtainable. The criterion or dependent measures to be used in such a design, the sources of variation to be controlled; and the dimensions of cost to be attended to are summarized in Table 3.2

A note of caution needs to be introduced at this point. The fact that practice oriented studies require the careful control of unwanted sources of variation in both treatment conditions and outcome measures has led educational researchers historically to undertake practice oriented research in a particular way. First, it has tended to be carried out within contexts that would "tolerate" such manipulation—for example, an instructor's own class or some limited segment of a student population. Second, it has tended to involve definitions of educational practices that were relatively narrow or limited in scope, for example, a one day or one week program of instruction, a particular grading practice, or a particular set of role relationships between instructor and student that are to be tried for a term. Both strategies have given rise to the problem of the results of such studies having limited utility—even within the context where they were done—and limited generalizability.

These are the problems that have plagued educational researchers for decades. Only now are we coming to understand what must be done to overcome them. Basically they are problems that have to do with the external validity of education experiments. Shulman speaks to such problems pointedly when he says:

> Researchers are caught in a bind. To maximize the *internal validity* of experiments, they develop carefully monitored settings within which they can govern their research. This has long been recognized as a necessity, but it is likely that the experimental tradition in American education over-emphasized the importance of reliability and precision at the expense of the characteristics affecting that other factor of equal importance (external validity) . . . It is not sufficient that the individuals studied as a sample are truly representative of that human population to which the results of a particular experiment will be inferred. Researchers must also ascertain that the experimental conditions can serve as a sample from which to make inferences to a population of external conditions of interest. That is, researchers must also attempt to *maximize the similarity* between the conditions in which they study behavior and those other conditions, whatever they may be, to which researchers may ultimately wish to make inferences. The similarity should hold between psychologically meaningful features of the settings, not merely between the manifest aspects of the two situations. (1970, p. 377)

If Shulman's analysis is correct, three fundamental shifts in the nature and focus of educational research must take place if it is to be responsive to the criteria of external validity: 1) a shift in emphasis from the sampling of people to the sampling of educational environments; 2) a shift from the

BEST COPY AVAILABLE

study of single, isolated variables to the study of the complex of variables that make up ongoing educational environments; and 3) a shift from piecemeal, unrelated, "one-shot" research efforts to studies that are articulated through time, theory, and problem focus. It is the contention of the research committee of the Consortium of CBE Centers that CBTE programs provide the kind of contexts where such shifts, for the first time, have some hope of coming about.

Having said this, it is important to point out that the conduct of long-term, multi-faceted, contextually valid experimental research lays heavy demands on an ongoing educational context, and unless the context is special in many ways, such demands will not be met. In the judgment of the committee, at least five conditions must prevail:

- *Persons responsible for the management and operation of the program must be inclined toward experimentation.* Commitment to empiricism and the desire to know must be dominant features of the context. Research must be viewed as an integral part of program operation, and as such viewed as a continuous, necessary, and desirable part of the program.
- *Persons responsible for the management and operation of the education program must view the program as subject to continuous change, and view as a major data source for its change a systematically designed program of research on its effectiveness.* When viewed in this way, research can have both the immediate and applied value that is needed for it to be supported by those responsible.
- *Data of a quality that will support trustworthy research must be collected as a normal part of program operation.* Accurate, reliable descriptions of program operations must be provided and accurate, reliable measures of learning outcomes must be recorded as a matter of course.
- *Sophisticated data management, storage, retrieval, and display capabilities must be available.* Data to be used for research, program operation, and program adaptation purposes must be routinely stored on computers and routinely retrieved in formats that support program related decision making and research.
- *There must be an advisory structure that ensures that the research that is pursued has value to persons in the program as well as to the profession at large.*

It is recognized that these five conditions represent an unusual array of features to find in any ongoing educational context today. It is also recognized that they are difficult conditions to bring about if they are not already in existence. It is the judgment of the committee, however, that CBTE programs represent the most hopeful development in the history of education for such conditions to be met.

Interinstitutional Studies

While most practice oriented studies will be carried out within a particular institution, at least initially, the continuing concern for the utility and gen-

BEST COPY AVAILABLE

generability of the results of such studies requires that they be replicated ultimately in a variety of educational contexts. To some extent, this requirement can be met by testing a particular practice in different programs, or with different populations of students within a single program. In the long run, however, this is not sufficient. For until it can be shown that a particular practice is successful in a number of institutions, it is not likely that it will have a great deal of acceptance in or utility to teacher education at large.

In carrying out interinstitutional studies on alternative practices, it will be necessary of course to control for the same sources of variation and to take the same measures of cost that are outlined in Table 3.2.

Part III: Basic Studies

Historically, research in education and teacher education has been disappointing (Gage, 1963; Smith, 1972; Travers, 1973). A host of explanations have been offered for this fact, including the argument that the world of education is simply too complex to yield to the strategies of science, but two explanations seem to have most authority. The first is the failure of education research to meet the requirements of external validity. The second is its lack of adequate dependent or outcome measures. No matter how powerful a research design, or how powerful a treatment condition, research results of significance will not be forthcoming unless the measure of effect is strong.

One of the most hopeful features of the competency based education and teacher education movements is that they are designed specifically to overcome the perennial problem of outcome measurement. If carried to successful completion, competency based education should yield new and powerful measures of learning outcomes in pupils, and competency based teacher education should yield new and powerful measures of teacher performance. If these measures are of the kind and quality anticipated, and the research community recognizes them as such and takes advantage of their availability, basic research in education and teacher education should profit immensely.

Two conditions must be met, however. If competency based education and teacher education programs are to become contexts for basic research. First, as implied in the above paragraph, the measures employed in the assessment of teaching competence must be of a quality that permits their use in research. That is, they must be valid, reliable, and sensitive. Second, experimental designs must be employed in the context of program operation with sufficient rigor that "causality" can be attributed to the experimental or treatment variables investigated. Both of these conditions are above and beyond the requirements of normal program operation, but both can be achieved if introduced with care and foresight. When both are met, much is to be gained, for basic research can be carried out as an adjunct to normal program operation at little added cost.

While it is possible to combine basic research with program operation by meeting these two conditions, it needs to be pointed out that considerable risk is involved in attempting such a venture. High quality measures, for example, are often difficult and costly to obtain. Also, requiring that pro-

[illegible]

BEST COPY AVAILABLE

learner characteristics, in their pursuit of laws governing learner environment relationships that are generalizable to teaching, experimental psychologists have not attended (or, behaviorism), to either of these classes of information. The assumption underlying the present effort is that by highlighting the instructional process, and by attending to it both considerably and empirically, it will be possible in time to develop a prescriptive science of instruction which in turn will permit an effective and efficient technology of instruction (1967) p. 66).

A Paradigm For Basic Research In Teaching

Without wishing to become embroiled in the paradigm-model theory debate (Kuhn(ed), 1962; Eagle, 1962; Snow, 1973), it is necessary to present the conceptual framework that underlies the proposals in the study. Whether the framework is a model, a paradigm, or a first cut at a G. de Finetti's process. What is clear are the directional implications of the framework for research on teaching. Since these directional implications are not evidently seen as ... defining the significant problems and methods of a research field for successive generations of practitioners' (Kuhn, 1962, p. 10), the framework is explicitly referred to as a paradigm.

The paradigm proposed as a guide to basic research on teaching reflects a point of view about teaching that is not inconsistent with how teaching is viewed generally, but is inconsistent with how most research on teaching has been done. The paradigm can also be described as a 'context dependent' paradigm, for it is based on the assumption that for teaching to be effective what a teacher does must always be adjusted to the nature of the learning outcome(s) desired, the nature of the learner(s) being taught, the nature of the setting within which learning-teaching is occurring, and whether a certain about how learning occurs. As such, effective teaching is seen as representing an articulation of a considerable number of variables, especially when a wide variety of learning outcomes are being pursued by a wide variety of learner types under a wide variety of learning conditions. The paradigm is portrayed schematically in Figure 2.2.

Two comments seem to be made about the paradigm. First, it has strong directional implications for research on teaching. It specifies the major classes of variables to be attended to in such research, and it specifies that these variables be treated in designs that accommodate them as interdependent or 'contextually dependent' variables.

Second, it lends perspective to the anxiety that many educators feel about the continued emphasis in CBTE on teacher behaviors and the continued press to carry out research that attempts to link teacher behavior to pupil outcomes (see articles by Rosner and Kay, McDonald, and Frandsen) in the January, 1973 issue of the *Phi Delta Kappan* as examples of the continued press for such research). Within the framework of the proposed paradigm, teacher behavior may be treated as either a dependent or an independent variable, and when it is treated as an independent variable it must be done so within a design that calls it to student characteristics, setting characteristics, class of learning outcome pursued, and the learning processes in-

BEST COPY AVAILABLE

Even though illustrative CBTE programs are now beginning to emerge, a host of problems remain before the kind of research and documentation called for in the present paper can be undertaken. Many such problems are referred to in the course of the paper; for example, finding CBTE programs that are willing to face comparison with other programs or that are willing to accommodate the demands of ongoing practice oriented and basic research studies. Other problems remain, however, and need at least to be identified. The more important of these would seem to be:

- Assuring the continued development and operation of competency based teacher education programs as demonstration and research contexts;
- Securing the funding level needed to carry out the research and documentation studies outlined in the present paper;
- Finding competency based teacher education programs that are at a level of development and possess sufficiently competent staff to carry out the kind of research and documentation programs outlined;
- Fixing responsibility for the long term coordination and management of the research and documentation program called for.

A number of these problems revolve around the availability of funds. This is obviously the case with respect to the research activities that have been proposed. Funding is also involved, however, in the continued development and operation of CBTE programs. Unless viable CBTE programs exist, most of what has been proposed is, in the parlance of the day, inoperative.

Some of these problems are not so much a matter of money as they are the availability of personnel who are competent to carry out the functions that will need to be carried out. In this regard, two broad sets of functions need to be performed—developing and maintaining a CBTE program that can in fact support the kind of research that has been called for; and actually carrying out the research and documentation activities.

Finally, a number of these problems are plainly and simply a matter of the National Institute of Education and the U. S. Office of Education agreeing to the importance of the kind of research and development effort that has been outlined; the probability of its short- and long-term contribution; the priorities of the proposed program as compared to the priorities of other programs that are being supported; the priorities of the various activities called for within the overall program; and the responsibilities to be assumed individually and collectively by the two agencies for the implementation, funding, and management of the program. In approaching the various problems and issues that have been identified, it is recommended that the last issue mentioned be attended to first, since decisions reached at that level will largely determine what happens in relation to all else.

Some Optimistic Projections

Assuming that the National Institute of Education and the U. S. Office of Education were able to reach the decision that the proposed program was worth supporting, that it had high priority in relation to other programs, that agreements were able to be reached in terms of shared responsibility

BEST COPY AVAILABLE

for supporting and managing the program, that a reasonably high level of funding was made available to the program, and that operational CBTE programs could be found that were willing to serve as research contexts, what could be expected to emerge from the program at the end of a 3-year period? at the end of 5 years? at the end of 10 years? The projections that follow attempt an answer to these questions. They assume that all lines of inquiry that have been recommended are to be supported.

Products And Conditions At
The End Of Three Years

DATA FROM POLICY ORIENTED STUDIES

- Preliminary data on costs and benefits of CBTE and non-CBTE programs
- Preliminary data on costs and benefits of alternative forms of CBTE programs
- Methodologies, settings, and personnel established to carry out long-term policy studies

DATA FROM PRACTICE ORIENTED STUDIES

- Preliminary data on the costs and benefits associated with selected instructional practices within particular CBTE contexts
- Methodologies, settings, and personnel established to carry out long-term studies of the costs and benefits associated with selected instructional practices within particular CBTE institutions
- Methodologies, settings, and personnel established to carry out cross-institution replications of cost and benefit studies on selected instructional practices

DATA FROM BASIC RESEARCH STUDIES

- First level taxonomic work completed in relation to the construct delineation studies, related literature reviews completed, design and measurement methodologies tested, and some replication data on constructs that have appeared promising in existing education and psychological literature
- Theoretical positions that link the constructs identified in the research paradigm stated at an initial level of detail, related literature reviews completed, design and analysis methodologies tested, and studies in progress that test the soundness of the theoretical structures that guide the assembling of variables

CASE STUDIES FROM THE DOCUMENTATION EFFORTS

- At least three case studies on selected CBTE institutions covering at least two years of development and implementation decisions
- At least six case studies on selected attributes of competency based teacher education that have 2-year history of development and implementation

BEST COPY AVAILABLE

Products And Conditions At The End of Five Years

DATA FROM POLICY ORIENTED STUDIES

- Sufficient data on costs and benefits of CBTE and non-CBTE programs that policy decisions can be made with reasonable confidence
- Sufficient data on costs and benefits of alternative forms of CBTE programs that policy decisions can be made with reasonable confidence
- Methodologies, settings, and personnel established to carry out whatever additional policy studies are needed

DATA FROM PRACTICE ORIENTED STUDIES

- Sufficient data on the short-term costs and benefits associated with selected instructional practices in particular CBTE contexts that program implementation decisions can be made with reasonable confidence
- Preliminary data on the long-term costs and benefits associated with selected instructional practices in particular CBTE contexts
- Initial replications across institutional settings of cost/benefit studies on selected instructional practices

DATA FROM BASIC RESEARCH STUDIES

- An initial set of empirically grounded constructs within each cluster of variables identified in the research paradigm that has been proposed
- Measurement methodologies firmly established for each construct within each cluster of variables in the paradigm
- The theoretical structures that link the constructs identified in the research paradigm reasonably well worked out and reasonably productive in terms of hypothesis generation
- Design and analysis methodologies firmly established for research on principles of teaching, given the constraints of the paradigm governing principle-related research
- A beginning collection of empirically based principles of instruction that are suggestive of the nature of the interactions between the variables listed in the paradigm on teaching

CASE STUDIES FROM THE DOCUMENTATION EFFORTS

- Completed case studies on three selected CBTE institutions (each case study would cover a 4-year period of time)
- Six additional case studies of selected attributes associated with competency based teacher education programs that have had at least a 2 year history of development and implementation (making a total of 12 such documentaries)
- At least three case studies covering a 2-year period of time on institutions that are serving as centers for research in CBTE

BEST COPY AVAILABLE

Products And Conditions At
The End Of Ten Years

DATA FROM POLICY ORIENTED STUDIES

- The completion of whatever cost/benefit studies are needed on CBTE as opposed to non-CBTE programs
- The completion of whatever cost/benefit studies are needed on alternative forms of CBTE programs

DATA FROM PRACTICE ORIENTED STUDIES

- Short-term cost and benefits studies essentially completed on selected instructional practices within particular CBTE contexts
- Sufficient data on long-term cost and benefits of selected instructional practices in particular CBTE contexts that program decisions can be made with reasonable confidence
- Sufficient cross-institutional replication data on cost-benefit studies of selected instructional practices that program decisions by a wide variety of institutions can be made with considerable confidence

DATA FROM BASIC RESEARCH STUDIES

- Second generation construct delineation research underway and the theoretical structures within which constructs rest under empirical test
- Second generation design and methodology in relation to construct delineation studies under development
- A sizeable number of principles of teaching established, and established with sufficient confidence that program design as well as online teaching is able to benefit from their application

CASE STUDIES FROM THE DOCUMENTATION EFFORTS

- All case studies on CBTE programs and program elements completed
- All case studies on CBTE programs as contexts for research completed

Figure 3.5 represents a translation of these projections into a single time/task chart.

A Plan Of Procedure

Any of a dozen strategies could be followed in implementing the program of research and documentation that has been proposed. Each would have advantages and disadvantages, and more than one would undoubtedly prove to be successful. Be that as it may, a number of recommendations are offered at this point to assist in its implementation. The recommendations assume NIE and OE agreement as to joint sponsorship of the proposed program and agreement as to the division of responsibilities therein.

Figure 3.3

A TIME/TASK CHART THAT IS BEING PROPOSED

(Solid lines represent periods of major research activity. Broken Lines represent beginning or closing activities)

BEST COPY AVAILABLE

BEST COPY AVAILABLE

1. Begin work on all aspects of the research proposed, even though there may not be enough money to do so on a large scale. All of the research efforts outlined need to be done; each will contribute to the other, and each will serve the needs of different audiences.

2. Take fully into account the program development work needed in order to carry out the research that has been proposed, and then see to it that the development and the research aspects of the overall program are closely coordinated. Without such attention to development there is danger that the research contexts on which the entire proposal is predicated will not be available.

3. Share the research and development costs across Federal agencies; for example, have NIE assume the costs of research and OE assume the costs of development and documentation.

4. Keep responsibility for managing the overall research program a Federal responsibility, rather than a state or non-Federal responsibility.

5. Assign the program on competitive proposals that are in response to RFP's.

6. Direct most of the funding for research to institutions that have CBTE programs that can function as Research and Development centers (funds should not be directed to existing Research and Development Centers unless they are linked to CBTE programs that meet the criteria for CBTE research centers).

7. Before letting research contracts to any institution, establish carefully the characteristics needed by an institution to carry out the various research and development activities that have been called for.

8. Implement a dissemination function at major proportions, and target the research and documentation results to be disseminated to the audiences for which they are particularly appropriate.

Obviously, the research program proposed is a massive undertaking. It is also an extremely complex one, and one that in the minds of many would be too slow to produce results or too grandiose ever to get started. In the judgment of the members of the Consortium of CBE Centers, however -- and it is a judgment that has been carefully weighed -- it is precisely this kind of research program that is needed in order for teacher education programs ever to make the kind of difference in teaching that they are supposed to make. The irony and hope is that by using CBTE programs as contexts for research, costs will be sufficiently reduced and the problem of the external validity of experiments will be sufficiently overcome so that the program that has been proposed can be carried out successfully.

BEST COPY AVAILABLE

PART II

ASSESSING PROFESSIONAL COMPETENCIES

4 Listening to Different Drummers: Evaluating Alternative Instructional Models

BEST COPY AVAILABLE

only important element in the creation of the environment. Setting becomes extremely important to both the learner and the teachers. In recent years in our work at the Agnes Russell School, we have created a variety of learning centers which support different kinds of teacher activity. In the skills center, for example, the teacher works with the assistance of a variety of diagnostic devices and a range of systems which can be used by learners to remedy particular kinds of skill difficulties. This teacher in the skills center is facilitated in the role of skill builder by these support materials. In addition, he is affected by the concentration of his efforts, which has enabled him to train himself in specific role-related skills. He can use programmed instructional methods and other systematic teaching methods with a far greater accuracy and efficiency than can some of the other teachers in some of the other settings. Similarly, the teacher in the Sciencing Center is supplied with a storehouse of equipment which students can use to replicate science experiments. The presence of that equipment, as well as the teacher's competence in using it, affects strongly the kinds of environments that can be created within that Center. Students can generate experiments, organize their equipment, and have places where they can keep the material for long periods of time while they're conducting a series of investigations.

The Film Center was supported with videotape equipment, motion picture cameras, still cameras, and a darkroom. The teacher thus supplied was able to create an environment in which children learned the skills of filming and also could engage in production of films and television recordings.

Not only were the physical settings important, but also the setting of the Centers in the community of the school. For example, the Film Center worked on a contract basis with small groups of children. The Skills Center, on the other hand, was responsible for the diagnosis of reading and arithmetic difficulties for every student in the school. Appointments at that Center were mandatory for all children. The Sciencing Center negotiated with teams of teachers and their children to generate units of activity, as well as contracting with small groups of children. The possibilities of teaching are greatly affected by the fact that these Centers relate to their clients by appointment and by negotiation, as are the competencies needed to operate in the Centers. Competence is enormously related to both the material and social aspects of setting. The settings in many schools are so primitive that probably only a few kinds of educational environments can be created in them. The secondary teacher, working in 45-minute periods in a departmentalized organization has brief, if regular, contact with children. The roles he is able to generate in that kind of setting contrast greatly with those of the teachers in the Centers alluded to above. We need to increase knowledge about how setting influences competence, including how it is possible to activate different kinds of learning environments in a variety of kinds of settings.

Teaching as a Boost to Learning: Self-Directed Learner Activity as the Baseline

When one conceives of teaching as the creation of learning environments, the emphasis moves toward the learner. If learning is seen as a product of

BEST COPY AVAILABLE

the learner's interaction with the environment, then it is primarily the learner's activity which produces the learning. A teacher as an agent can help shape the environment and modulate it to the requirements of the individual child, but in the final analysis, there is truth in the old cliché that learning is done by the learner. Teaching can but facilitate learning, even in the most directive of the Models of Teaching. Because of this, the effects of a teaching competence—the extent to which it contributes to learner growth—needs always to be measured against what the learner might do for himself. The results of this can be somewhat surprising. David Hunt and I and our associates have conducted a series of studies in Toronto in which Marsha Weil and I have created environments styled after several models of teaching.[7] We have exposed to those environments students of varying personality structure. We have compared the substantive learning (recall of information and the integration of subject matter) of students exposed to those environments with that of children who read the material for themselves and conducted their own analysis of it. Even after five or six hours of teaching we failed to raise recall scores above those of the "do-it-themselves" control group. After as little as one hour, however, we affected the numbers and types of concepts the children formed—the integration of material. In other words, we as teachers and the environments created by the models of teaching have a much greater impact on the learner with respect to "higher order" outcomes.

Studies of the effects of teaching should be measured against the effects of self-directed learner activity. (The same principle applies in teacher education; training strategies should be measured against teachers' self-training.)

The Multidimensionality of Educational Environments

Educational environments consist of a variety of cognitive and social tasks set in a social context and a support system of materials. The learner thus is exposed to a social system, some kind of sequence of social tasks, sets of cognitive tasks, learning materials, and the personalities of the individuals in the environment. The study of the effects of instruction is the study of the results of being exposed to a many-sided environment. We can expect multidimensional environments to have several kinds of effects resulting from a combination of the substances into which one inquires, the cognitive and social tasks, and the social relations.

An instructional model or model of teaching is a pattern for the design of such an environment. In practice, inquiry into the effects of models of teaching is complicated by the fact that so many instructional strategies do not explicitly attend to a number of the important dimensions of the environment that they will inevitably create. We have to flush them out by careful conceptualizing and by empirical work if we are to know what we are really studying—what the educational environment really is. For example, the introduction into a school of automated instructional systems with individual study carrels influences the social as well as the instructional dimensions of the environment, although normally only the instructional dimension has been attended to. We have found, for example, that many

TABLE 4.1

SCRIPTING A MODEL - CONCEPT LEARNING

Task	Required Learner Skill	Involvement Options	Hypothesized Model-relevant Outcome	Content Outcome Boosted
1. Concept Attainment exercise	Discriminate attributes Infer principle Analyze strategy	Reception-Selection Sharing Strategies (Move) Student Leadership	Concept-attribute Concept of classification	Integration
2. Enumerate data	Recall Select by criterion e.g. heterogeneity triviality abstract-concreteness	Individual - Group (Number) Student Leadership	Enumerate data by criterion	Recall
3. Select data	Develop criteria Apply criteria	Group (Number)	Apply criteria	Recall
4. Group data: Initial pass (Includes Label)	Generate Groups	Individual - Group - Class (Number)	Group basis of common attributes	Integration
5. Share - classify - justify - predict labels	Identify common elements of grouped data Generate multidimensional possibilities	Student - talk	Identify principles of classification Concept of classification	Integration
6. Regroup - label data of some group members	Regroup - relabel data	Student - talk	Concept of classification	Integration
7. Multidimensional grouping	Generate multidimensional possibilities	Number	Multiple classification	Integration

BEST COPY AVAILABLE

8. Share, classify, justify	Identify principles of multidimensional grouping	Student-talk	Concept of multiple class	Integration
9. What is a label? (Share labels)	Analyze labels	Student-talk Leadership	Nature of Label	Integration
10. Alternative Labels (Generate and share)	Generate labels which refer to common attributes Compare labels against criteria of appropriateness	Student-talk	Criteria for adequacy of labels	Integration
11. Group labels	Generate concepts of concepts	Student-talk Number Leadership	Group Abstractions	Integration
12. Share Groups	Analyze concepts of concepts	Student-talk Number Leadership	Criterion for grouping abstractions	Integration
13. Label grouped labels and share	Generate labels	Student-talk Number Leadership	Ability to use labels for grouped labels	Integration
14. Characterization Task (Newspaper Story)	Use labels to operate on data	Student-talk Number Leadership	Application of labels in descriptive tasks	Integration

BEST COPY AVAILABLE

BEST COPY AVAILABLE

students do not have to learn, in such a setting for purely social reasons.

The closer we get to the study of learning environments, the more differentiated are their elements. In Table 4.1 there appears a task analysis of some of the components of a concept-learning model of teaching.

In the second column are the cognitive tasks which are presented to the learners. In the third column are the involvement options—the options that a teacher has with respect to the social system which will be utilized. In the first column on the left is an identification of the kinds of learning skills required to respond to the different tasks. This particular sequence of tasks can be carried out in a little over an hour with junior high school aged children. It provides some idea of the complexity of the environment that is created within a simple model used for purely instructional purposes.

In order to validate any given teaching competence, we need to analyze the dimensions of the environment thoroughly and to study them with respect to their effects on the learner. Only in highly controlled studies can we vary small aspects of the dimensions of the environments, but it is necessary to employ highly controlled studies—to conceptualize the environment very clearly and to manipulate variables systematically—if we are to obtain any reasonable knowledge about the effects of teaching on learners.

The Multiple Outcomes of Educational Environments

The learner in a multidimensional environment always produces a variety of kinds of outcomes.7 There are substantive outcomes, such as the recall of information and the development of concepts for integrating it; there are attitudinal outcomes, such as a liking for the material, the method, and the teacher. There is the learning of social and cognitive skills. All of these need to be taken into account. Specific instructional models may promote particular kinds of learning (what Marsha Weil calls model-relevant outcomes), but many models are likely to share educational outcomes, and may boast the same outcomes about equality. We need to test hypotheses about the existence of model-relevant outcomes and about cross-model effects. It is more likely that particular high-order and effective outcomes will be boosted by specific models than that paired-associate learning will. Exposure to material is rarely to be such a powerful determinant of sheer paired associate learning by a highly-motivated learner that it is unlikely that very many models will stand out. The integration of information, however, and the ability to apply information to new situations may vary much more.

I believe that the transfer of learning should receive the greatest emphasis in the evaluation of instructional strategies and hence, teaching competencies should be validated much more by the effect on transfer than by substantive outcomes. Consider, for example, the training of teachers. When we try to evaluate the effectiveness of strategies for training teachers, all of the important criteria are ultimately bound up in problems of transfer. If we teach an instructional strategy to a teacher, our only way of evaluating his performance is to learn what effects it has on children as the teacher transfers it to new situations.

BEST COPY AVAILABLE

It is equally important to know what criterion transfer from any given instructional encounter. How do they carry the information of a new unit-tion, how do they use the concepts they have learned, how do they apply the principles that are being developed, how do the attitudes they have formed and the social relations they have developed affect their subsequent behaviour?' Simply to evaluate the performance of a model of teaching, or a given teacher who uses it in terms of what is observed from the encounter itself is to concentrate on the trivial and ignore the important. Probably the most important learning outcomes of models of teaching are those which are embedded in the patterns of the models themselves and affect responsiveness to future situations. For example, one outcome of a counseling environment is to equip the learner to respond to nondirective encouragements. In the case of learning strategies centered on the models of inquiry of the disciplines, the most important outcomes are those which have to do with the application of those models to new situations. The model of learning embedded in the model of teaching is the critical transferable content. In the studies Dave Hunt and I have done recently, we have studied the process by which children acquire a model of teaching to the point where they can actually carry it out by themselves. We have exposed learners with different personality structures to training programs until they are able to behave in the mode of the model independently, without the assistance of the teacher. We have been very much surprised by the relative success with which learners pick up the subroutine of a model. What has not surprised us so much, but which complicates our lives considerably, is the difference in manner by children of different personality types. The more rigid children, for example, tend to apply a model more or less as the teacher uses it, following it almost as a formula, where the more complex children tend to transform a model and apply its heuristics, rather than the formula of patterns that they have been exposed to.

The study of transfer is enormously complicated by the fact that the learner transforms what he learns as he applies it in new situations. This plagues the study of transfer in teacher education. A teaching skill, which stands out clearly in a specific type of microteaching laboratory exercise, may appear in quite a different form and even be almost buried in the events of the classroom. For example, several of my associates and I for several years studied pre-trained teachers' sensitivity to learners in simulated teaching situations. In the laboratory settings the teacher was given a teaching task and presented with one or more learners (actually role players). The role players emitted cues which represented a point of view which conflicted with the teaching task. Until the teacher read the cues, attained the concept of the learner, and modified his behavior to build a bridge to the teaching task, he was unsuccessful in the exercise. We found striking individual differences in teachers' ability to "read" the learner and modulate to him, and in the styles which were employed to accomplish this modulation. We also found great individual differences in teachers' ability to acquire the "skills" of sensitivity. When we followed teachers into the classroom, however, we found no reliable way to detect those skills in the classroom the teacher was trying to read and modulate to many children, sometimes simultaneously. He transformed his skills enormously

BEST COPY AVAILABLE

in order to relate to the target group of children. Sometimes we observe [illegible] of the lessons we were after, but in general we were unable to develop an instrument and our system and which to harvest it, we plan to judge its effects on children.

Proposed Focus on Learning Outcomes. Presently, my associates and I are exploring the following types of learning outcomes as we explore the effects of different models.

The subsidiaries of the interventions. We explore the acquisition of information, general associate learnings, concepts, and principles that integrate the material which is exposed to the learner during any given period of time. Multiple choice items are used at the general associate learning level and interviews or interviews are used to collect data which is content analyzed to determine the kinds of concepts and memos which the learner develops in the course of the inquiry.

Model relevant concepts or skills. We have been exploring the development of ways to identify the particular skills and concepts which are general to particular kinds of models. The synthetics model of teaching, for example, attempts to teach children to think metaphorically. If they explore a given subject matter in metaphoric mode, they are acquire information and form concepts, but they also learn the concept of metaphor and how to use it. This is a model-relevant concept, and the ability to use it is a model-relevant skill. Group investigation requires learners to work together to define objectives and to analyze one another's perceptions in a social situation. The ability to work with others to identify objectives becomes, therefore, a model-relevant skill for group investigation. Concept learning teaches children to enumerate, classify, and label data. The concept of classification, the concept of labeling, and the ability to classify and label are therefore model-relevant outcomes for that model.

Transfer or subsidiaries of outcomes. This outcome involves the ability of the learner to apply what he learns to a new situation. In the synthetic we teach children how to analyze a culture we could then investigate their ability to analyze a second one, or perhaps compare the second analysis with the first.

Acquisition of the models of teaching. I referred to this earlier when I described our attempt to determine what it takes to help a student learn to use the cognitive and social skills of a model of teaching to guide his own inquiry. A case example is when we use inductive methods with children and then study their ability to carry out inductive inquiry independent of us. Situational tasks are most appropriate for investigation of this kind of outcome.

Attitudes toward self. Inventories can be used to investigate how the learner views himself when interacting with any environment. Inventories can also be used to study the following affective outcomes:

Attitudes toward material. (How does the student feel the curriculum and learning material he encounter item in an environment?).

Attitudes toward the method. (How does he feel about the methods of teaching he experiences?).

Attitudes toward the teacher. (How does he feel about the training agent?).

These are described in Tuttle Two.

BEST COPY AVAILABLE

5 Task Force on Relating Teacher Competence to Pupil Outcomes

Donald M. Medley
David Krathwohl

A substantial portion of the time allotted to the task force was spent in small groups whose mandate was to formulate questions for discussion when the entire task force reassembled. Since not all of these questions were dealt with by the larger group, a summary is appended to this synthesis.

The issue that seemed to receive the most attention had to do with whether it is defensible to attempt to define and assess teacher competence in terms of pupil outcomes. The wording of the charge given to the task force—to deal with the problem of relating teacher competence to pupil outcomes—implies that teacher competence is (1) a different entity from pupil outcomes, and (2) that the relationship between the two is an empirical question. While the point was disputed, most of the group concurred in the opinion that the teacher should be held accountable for competence rather than outcomes.

An analogy often cited was the medical one: it was pointed out that when a patient dies, the physician is not held responsible so long as he has correctly used the best known methods for diagnosing and treating the patient. It is generally realized that the state of medical research is not sufficiently advanced that sure cures exist for many diseases. Failures tend to be blamed on incomplete research data rather than on the incompetence of the individual physician. When the practitioner's competence is called into question, his defense is based on his knowledge of and ability to apply the best known methods, and not on outcomes.

Some question was raised about the fairness of the evaluation procedures used in the program described by Del Schalock, in which students are re-

BEST COPY AVAILABLE

quired to prove their ability to produce pupil learning before they may be certified as competent. As one member phrased it, "If the student uses the correct procedure according to his supervisor, and the pupils fail to learn—who should be held accountable for the failure—the student or the supervisor?" Or, to paraphrase another comment, is it ethical to demand that a teacher demonstrate more knowledge of the teaching process than those who purport to be training him possess?

The state of research in teacher effectiveness—that is, research relating teacher behavior (process) to pupil outcomes (product)—is far less advanced, everyone agreed—than research in medical science—that is, research relating the behavior of physicians to patient recovery.

The defense of the Schalock program was that the learning outcomes to be produced were determined by negotiation between teacher and supervisor, and were restricted to what would be reasonable to achieve with the group of pupils in question. This is an important point, and goes a long way toward removing the objections cited above by ensuring (to a degree) that the student is not confronted with an unreasonable task.

Serious questions were raised about the validity and reliability of such a procedure. The reliability is, of course, unknown, and inaccessible to study since each performance is based on different criteria as well as on different pupils, content, etc.; further, not only the evaluation but the definition of the criteria are highly subjective. Decisions as important as whether or not a student is to be admitted to the practice of a profession should be based on measures of high objectivity and reliability, as well as validity.

Questions regarding validity took two forms. The desirability of certifying teachers on the basis of ability to produce only short-term gains in limited content areas and with only one type of class was debated. The fact that the student is required only to produce what his supervisor considers a reasonable amount of learning was seen as tending to stifle progress by requiring the new teacher to be no more effective than the old.

All of this discussion seemed to serve the purpose of sharpening a distinction between the activities of research and evaluation. When we are evaluating teachers individually, we need much greater reliability (and practicality) than we need when we are doing research in teacher effectiveness.

The objective of research is to identify teacher competencies that are related to pupil outcomes in a probabilistic sense, that is, teacher behaviors which are most likely to produce pupil learning. Research studies are statistical in nature and deal with groups of teachers; if we can show, for example, that the mean gain in a group which behaves in one way is greater than the mean gain in a group which behaves in another way, we have evidence of the greater effectiveness of one way of behaving. Reliable mean differences can (fortunately for us) be demonstrated with measures whose reliability for discriminating between individuals is low.

Measures of pupil outcomes are highly unstable; the principle source of errors of measurement are differences between pupils used one time and those used another time with the same teacher. If the average correlation between gains measures based on different groups taught by the same teacher is around .3 (as Rosenshine reports), then it would be necessary to

BEST COPY AVAILABLE

average gains in seven groups to get a measure whose reliability is .75. Assuming that we are interested in gains achieved over a one-week interval, this means it takes some seven weeks to measure teacher effectiveness. If we are interested in gains over one year, it takes seven years! The impracticality of using mean pupil gains as a basis for teacher *evaluation*—even if one is willing to use very short term gains indeed—is obvious. And yet such measures can be used in *research* studies to identify competencies the possession of which will increase a teacher's effectiveness.

If a teacher is evaluated according to the competencies he possesses, adequate reliability can be achieved with a reasonable expenditure of time and effort. And when research begins to identify new ways of being effective, we can begin to certify teachers who are more competent—and more effective—than the ones we certify today.

When planning a study in this area, it is important to design a study with sufficient power to detect relationships that are none too robust; in particular, sample sizes of at least 50 to 100 teachers are needed. Data should be collected that provide information, not only on the teacher behaviors and the pupil outcomes which are the primary concern, but also about variables likely to interact with them. Pupil characteristics (especially socioeconomic status and related characteristics) and type of curriculum seem particularly important. It is dangerous to assume linearity. Recent research strongly indicates that in many cases, the regression of pupil growth on teacher behavior is curvilinear with optimum effectiveness achieved by teachers whose behavior is somewhere in the middle of the observed range of behavior. Outcome measures should reflect more than one type of pupil learning, since it is likely that different patterns of teacher behavior will be most effective in facilitating different kinds of learning.

To sum up, the general attitude of the group was positive—the consensus was that research in teacher effectiveness can increase our ability to train competent teachers in performance-based programs by identifying the nature of teacher competence. How rapidly these programs improve depends to a great extent on how extensive a research program we are able to mount and sustain.

Relating Teacher Competence to Pupil Outcomes Task Force

A. An area of concern to almost all members of the group had to do with the specification of the outcomes to which competencies are to be related. Questions raised include the following:

   1. What philosophical base should be used in determining the desired outcomes?

   2. How may parents (and other community groups) gain access to the establishment of desired pupil outcomes?

BEST COPY AVAILABLE

3. How will political constraints and decisions bread on them -- which may or may not be compatible with decisions about desired outcomes -- be dealt with?

4. Are there some competencies which relate positively to some outcomes and negatively to others?

5. Should long-term outcomes be used, and if so, for what purposes?

6. How can the public be educated to expand its concept of outcomes to include affective dimensions?

B. A second area of concern had to do with interactions between teacher and learner characteristics (and other extraneous variables) and their effect on relationships between competence and outcomes.

1. What variables need to be taken into account in identifying pupil learning styles, teacher learning styles, etc.?

2. Does the interaction between these two vary for such types of outcomes as self-concept, student attitudes, and mastery of cognitive skills?

3. Are there any generic teaching skills whose effects are invariant across such variables as these?

C. A number of researchable questions were raised, incuding the following:

1. What kinds of teacher competencies produce what kinds of outcomes with what types of pupils?

2. How are these relationships affected by entering characteristics of pupils and curriculum variables, and how do they interact with teacher behaviors?

3. Is it possible to isolate the effects of teacher behavior from other environmental factors?

4. What are some of these other variables?

BEST COPY AVAILABLE

3. Is there enough known correlation between competencies and outcomes to justify certification on the basis of competencies?

4. Which kinds of outcomes are most reliably measured at the pre-service level? at the materials level?

7. Are there competencies which once acquired are never lost? Are there competencies which need continual reinforcement and practice to be maintained?

D. Miscellaneous Questions

1. How can effects of setting and teacher role be separated?

2. What variables are involved in research relating teaching and learning styles?

3. Should the emphasis be on relating outcomes to competencies as such, or in identifying effective teachers and determining what competencies they possess?

Appendix

Summary of Questions Raised by Task Force A

1. Using the regression model, couldn't you regress the teacher behaviors on performance of the outliers?

2. How to overcome the resistance to judge certification eligibility based on pupil learning -- uncontrollable conditions, differences in learners, home background

3. The "invasion of privacy" issue which will emerge when we use data of practicing teachers: effect on career, certification, salary, unless all in population are included.

4. Are there ways of relating change in pupil performance to causal factors without experimental designs?

5. What about styles of teaching vs. child learning styles?

6. What are the act categories under consideration?

7. What concepts of curricular structure and instructional structure are being assumed?

BEST COPY AVAILABLE

6. Is the teacher assumed to be performing in his individual work under certain practical principles and ethical habits?

6 A Practical Approach to a Complex Problem

Part 1

George Gustafson

BEST COPY AVAILABLE

Part 2

Marjorie Powell

Teacher training institutions admit students who meet entry criteria and then move the students through the process, their teacher education program, designed to prepare students to teach in the public schools. The California Commission for Teacher Preparation and Licensing requires that institutions describe the program which they will implement. The Commission grants approval for a period of time to the program as described by the institution, and is establishing a procedure to insure that programs are implemented as described, or that modifications are documented.

While these steps involve the person entering the program and the process by which the person is educated, they do not yet look at the product, the teacher teaching within the school classroom. The important question is still unanswered. Will the students who move through the program as described and implemented by the institution be successful teachers when they have received their certificates from the Commission and have been hired to teach? What skills are necessary to be a successful teacher?

In order to look for the answer to the question of what is a successful teacher, the Commission has undertaken a research study, the Beginning Teacher Evaluation Study, with funding from the National Institute of Education. The purpose of the multi-year study is to identify teacher behaviors which are related to student achievement in reading and mathematics.

We have limited the focus of the study to reading and mathematics because of the necessity of confining the research to manageable proportions. Such a study has a way of expanding until the entire universe is included and the impact of the study is diffused to the point where nothing is learned. However, we do recognize that the elementary school curriculum contains more than the basic skills of reading and mathematics, and we hope that the work of this study will be expanded in the future to incorporate more areas of the elementary program.

The first year of the multi-year study, Phase I, was devoted to planning. After the development of an initial design for the study, the Commission sponsored a Research Design Conference in February, 1973. The conference participants, educational researchers, teacher educators, school administrators, and subject matter specialists, reviewed the research design and made a series of recommendations. Based on many of these recommendations, a design for the second year of the study was finalized.

The second year, Phase II of the study, focuses upon the development of instruments and the generation of research hypotheses. Educational Testing Service (ETS) was selected to conduct the work of Phase II during the 1973-1974 school year. Experienced teachers in grades 2 and 5 form the sample for Phase II, the first year of field work. School districts were asked to participate, then experienced teachers were contacted, provided with information about the study, and asked to volunteer.

During the planning for Phase II, a variety of procedures was discussed for identifying the teachers to be included in the sample. We considered a

BEST COPY AVAILABLE

process of nomination by all of the teachers in the school or by the principal. However, even when the criteria to be used to make the nominations are clearly specified, in many instances nominations are based on factors other than the specified criteria. We finally concluded that we wanted to study experienced teachers in the variety of instructional situations, without using nominations to identify teachers to be included in the study.

The Phase II sample of 97 experienced teachers in eight districts included teachers in urban, suburban, and rural districts teaching students from low and middle income areas. Some teachers are in self-contained classrooms while others teach in open space programs. Some have multi-grade groups of students, some are departmentalized, teaching only one or two subject areas. Several teams of teachers volunteered and are included in the work this year. The majority of the teachers have some form of teaching assistance -- an aide, parent volunteers, or older students who tutor specific children.

We are particularly pleased that teachers from so many instructional settings are involved in the sample during the first year of field work. We hope to identify teacher behaviors which are important in a variety of instructional settings and must insure that the measurement instruments which we use to identify and measure teacher behaviors are appropriate for, and applicable to, a variety of instructional settings.

While the major focus of the study is on newly certificated teachers, a model must first be built, based on behaviors exhibited by, and related to student achievement in the classrooms of experienced teachers. For this reason, experienced teachers will be studied during Phase II and will be the major part of the sample in the work next year.

Instrument development during Phase II has included the development of student achievement tests in reading and mathematics, specifically designed to measure student achievement in reading decoding, comprehension, application of skills and attitude; mathematics computational skills, concepts, application of skills, and attitudes.

Fall data collection was conducted during November, 1973. Student achievement will be measured again in the spring to provide a measure of student growth during the school year. Most of the student instruments have been developed specifically for this study, so normative information will not be available.

Instruments have also been developed to measure teacher knowledge of the teaching of reading and mathematics, and teacher verbal ability, or the ability of the teacher to process information. The latter instrument included subtests of vocabulary, memory, classification skills, inductive and deductive reasoning, and creativity. A film test is being developed to assess the skills of teachers to diagnose the reading problems of students and the range of alternatives which the teacher would prescribe to meet the identified need(s) of the students. The test is being developed because we recognized the importance of the skills of diagnosis and prescription, but we were concerned that we might not observe enough instances of diagnosis or prescription to adequately measure the skills of the teachers during the somewhat limited observation period in the classrooms.

Observation procedures and two instruments have been developed by ETS

BEST COPY AVAILABLE

specifically for the multiple classroom observations within this study. The observations are occurring from February through April of the 1973-1974 school year. Differing numbers of observations are being done using the two observation instruments. The teachers indicated that they would prefer as many observations as possible, since they did not foresee that a few observations would provide enough information about important teacher behaviors. The differing observation patterns will also allow for the analysis of the reliability of the instruments and the observers, and a determination of the varying types of information obtained through the use of each instrument when the same classroom is observed at the same time by observers using different instruments.

Teachers are also being videotaped, so that further analysis of teacher behavior can be completed. The teachers will have an opportunity to view their own videotapes after the taping is completed, providing the teachers with an opportunity to see themselves in instructional situations. In addition, information about both teacher and student background has been collected. A series of work diaries has been developed to collect information about the instructional program within the classroom throughout the school year and the particular instructional events during the specified weeks of the year. Teachers work within an instructional setting, with resources available to them, and with specific goals for students. The work diaries are designed to obtain data concerning the instructional resources and goals.

The school climate is being measured through questionnaires for teachers and principals, and an interview with the principals. All school instruction occurs within the setting established by the persons involved in the school: the principal, the teachers, and the students. This environment, or climate, has an impact upon the instructional situation and therefore must be assessed in order to adequately determine the relationship between teacher behaviors and student achievement.

The consistency of teacher behavior throughout the school year will be determined through the consistency substudy, which involves the videotaping of a sample of teachers three times during the school year, and the analysis of the videotapes in terms of the consistency of teacher behaviors across the three videotapes.

The stability of teacher performance, as measured by student achievement, will be assessed through the collection of historical test data for students in the classrooms of the sample teachers in previous years. This task is especially difficult due to the differences in testing programs across the eight participating school districts.

The dependent variable in the Beginning Teacher Evaluation Study is student performance, a measure of growth from fall to spring. The independent variables are teacher performance (behaviors), teacher characteristics (years of experience, etc.), student characteristics (age, language background, etc.), and school characteristics (size, special programs, etc.).

Phase III will be designed to refine instruments and to test the research hypotheses which are developed in Phase II, to provide empirical data concerning the relationship between teacher behaviors and student achievement in reading and mathematics. During Phase III, beginning and experi-

elected teachers will be included in the sample of teachers. To test the hypotheses with both types of teachers.

Phase III is planned to cover three years, with the first year devoted to additional development and field testing of instruments and research hypotheses. A series of special studies are planned for the first year. All of this work will lead into the second year of Phase III, during which a large-scale field study will provide a test of the research hypotheses. The third year will focus upon a review of the implications of the study for teacher certification and teacher preparation.

To review these implications and to obtain information about training procedures, a series of small studies will be conducted in which teachers will be trained to utilize the behaviors identified during Phases II and III. These studies will provide both a test of the relationship between the teacher behaviors and student achievement, and information about one or more methods of preparing teachers to exhibit these behaviors.

While I am confident that the study, over several years, will identify important teacher behaviors and will provide information about procedures for training persons to implement the behaviors, I would be unrealistic if I expected that the study would result in a final list of behaviors which could be defined as vital for all teachers in all teaching situations. Rather, I anticipate that we will identify some behaviors which are important to all teaching situations and other behaviors which are important to specific instructional settings or the teaching of specific skills. I am convinced further that institutions will learn information about the one or more methods of teaching students to exhibit the identified skills.

7 Classroom Observation

BEST COPY AVAILABLE

3. Interpretation - identify similarities or differences, relate supporting evidence to a generalization, or carry out a specified operation;
4. Application - bring together and use previously learned material such as deciding what mathematical operation is appropriate to solve a problem;
5. Analysis - applying the rules of logic to the analysis of a problem, or inferring feelings or motives;
6. Synthesis - bringing ideas together, but reorganizing or changing them in such a way as to produce something new;
7. Evaluation - establishing a set of criteria, and then evaluating.

The original instrument was intended for all grade levels, but since a performance which is routine and low level when well learned may be high level while being learned—reading, for example—a new version of the instrument was developed for kindergarten or first grade.[19]

Although it is sometimes believed that the higher levels are more appropriate for older pupils than for younger, experience indicates that this is not true. The difference, of course, is that simpler materials and concepts are involved at lower grade levels. The development of a Piagetian concept such as conservation would fall at the level of synthesis, and the discussion that accompanies a story or a reading lesson may deal with questions such as "What else might Jimmy have done?" (synthesis), or "Would it have been better if Jimmy had done something different? Why?" (evaluation).

Reciprocal Category System (RCS). This is an extension of the Flanders System of Interaction Analysis, which is one of the most widely used instruments and whose well-known 10 × 10 matrix first permitted analysis of sequences.[20]

In the modification by Ober, Wood, and Roberts, the seven teacher categories were expanded to nine.[21] Teacher lecture is divided into that which is responsive to pupils, and that which is teacher initiated; the category of teacher criticism is divided into a category for correction without criticism, and one with criticism. Category 10 remains silence and confusion as before. Each of the categories is redefined, however, so that it can be used for pupil talk as well as for teacher talk. That is, teacher amplification of a pupil's idea is coded as a 3; a pupil amplification is a 13. Each category is changed from a teacher category to a pupil category by adding a "1" as the first digit. The observer, then, learns nine categories as he did with the Flanders System but has 18 to work with and, as a consequence, the same variety of pupil talk is recorded as teacher talk. Practically speaking, this modification offers greater richness of data at little increase in the complexity of the observer's task. A strength of the matrix procedure initiated in Flanders' work is the capturing, one step at a time, of the sequence of occurrences by the way the categories are tabulated. An interesting feature of the RCS is that it can be broken down into four submatrices: teacher-teacher talk, teacher-pupil talk, pupil-teacher talk, and pupil-pupil talk.

Observation Schedule and Record, Form 5, Verbal (OScAR5v). This is the current revision of an instrument which has a history beginning in 1958.[22] Without question, the various forms of OScAR and the Flanders System of Interaction Analysis have been the two instruments used more than any

BEST COPY AVAILABLE

others by researchers.

The current revision of OScAR differs sharply from the original, and codes interactions as sequences. It uses 18 categories—four for pupil utterances and 14 for teacher utterances. Since six of the teacher categories are dual-purpose, however, there are, in effect, 20 for teacher behavior. The instrument provides for recording interactions by pairs of items which record either pupil utterance and teacher evaluation or response, or teacher question and the teacher's response to the pupil's response.

The fact that both entry to and exit from the interchange may vary means that the number of combinations—the number of specific interchanges—is large, and although the primary focus of the instrument is substantive, social-emotional behavior in management is also recorded.

One aspect of the instrument which seems likely to be important in a variety of settings and with a variety of objectives in PBTE is the ingenious system for recording interchanges—which not only records a teacher behavior, but also its antecedent. It seems likely that PBTE could use such a procedure for coding sequence with a variety of kinds of content.

Personal Record of School Experience (PROSE). This instrument is a system for recording the behavior and experience of individual young children.[23] The method used is time sampling. For a "cycle" of five observations, on signal from a timer, every 30 seconds the observer records the activity of the child at that moment. A series of 11 "words" are used to describe the activity; words 1-3 describe the adult-child contacts, with 1 recording the activity of the adult; words 4 and 5 record child-child interactions, with 4 the target child and 5 the child with whom he is interacting; word 6 records the level of the activity of the target child in the previous words, in terms of physical contact, materials, or whether verbal or nonverbal communication occurred; word 7 contrasts the sex and ethnic group of the target child with the other person; words 8, 9, 10, and 11 describe the appropriateness of the activity, the nature of the physical activity (movement), the nature of other activity (work, fantasy, etc.), and whether the activity was structured by a teacher or by a child.

After five observations (a "cycle") have been coded, the observer records the child's context in terms of the general activities of the classroom as a whole, other adults, and the instructional content.

Classroom Observation Instrument (COI). This instrument is the one which Stanford Research Institute has used to represent the major programs of planned variation in Headstart and Follow Through.[24] The instrument provides for a physical description of the classroom (recorded at the end of the day) in which numbers of adults and children are recorded, along with physical aspects such as noise level, lighting, furniture arrangement, and whether pupils select work groups. The observation begins with a Classroom Check List, which records the grouping patterns, materials in use, and adult responsibility. The Five-Minute Observation records four items of information for each of 76 frames: Who, To Whom, How, and What. The first two identify the initiator and receiver; What includes such items as command or request, open-end question, response, or instruction; How includes such items as nonverbal, motion, happy, or positive. The information recorded in each frame is summarized as a *sentence* which describes an event, and

BEST COPY AVAILABLE

provision is made for simultaneous or for recurring events.

Since there are thousands of possible combinations for a sentence, it is necessary to combine them into variables for further analysis.

Other Systems. There are hundreds of systems which have been developed, and many are reviewed by Rosenshine and Furst.[25] The most extensive collection of instruments, with annotations, is the series of volumes by Simon and Boyer.[26]

A Final Comment

When one reviews the major points of this paper, the problems in classroom observation seem overwhelming. We have often approached teaching-learning as a simple, linear relationship while it is neither simple nor linear. The statistical problems are great; the problem of measuring outcomes for pupils is complex; and the measurement of classroom behavior is even more difficult.

Yet there is no alternative to facing these problems, because the most important part of schools is what is happening in the classrooms.

FOOTNOTES

[1] B. Rosenshine, *Teaching Behaviors and Student Achievement*, (London, England: International Association for the Evaluation of Educational Achievement, 1971)

[2] R. W. Heath and M. A. Nielson, "The Myth of Performance Based Teacher Education," paper presented at the American Educational Research Association (New Orleans, Louisiana: February, 1973), 16.

[3] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. (New York: Academic Press, 1969)

[4] R. S. Soar, "Optimum Teacher-Pupil Interaction for Pupil Growth," *Educational Leadership Research Supplement 1*, (1968), 275-280.

I. J. Gordon, ed., *Early Childhood Education*, (Chicago, Illinois: National Society for the Study of Education, 1972). "An Empirical Analysis of Selected Follow Through Programs. An Example of a Process Approach to Evaluation," by R. S. and R. M. Soar, 229-259.

R. S. Soar and R. M. Soar, *Classroom Behavior, Pupil Characteristics, and Pupil Growth for the School Year and for the Summer*. (Gainesville, Florida: University of Florida, grant no. 5RO1MH15891, and 5RO1MH15626, Washington, D.C.: National Institute for Mental Health, Department HEW, 1973)

[5] Ibid.

[6] R. S. Soar, "Optimum Teacher-Pupil Interaction for Pupil Growth," 275-280.

R. S. and R. M. Soar, "An Empirical Analysis of Selected Follow Through Programs," 1972, 229-259.

R. S. Soar and R. M. Soar, *Classroom Behavior, Pupil Characteristics, and Pupil Growth for the School Year and for the Summer*, 1973.

[7] J. R. Squire, ed. *A New Look at Progressive Education*, (Washington, D.C., Association for Supervision and Curriculum Development, 1972). R. S. Soar, "Teacher Pupil Interaction," Chapter V, 166-204.

[8] R. S. Soar, *An Integrative Approach to Classroom Learning*, NIMH Project Numbers 5-R1MH01096 to the University of South Carolina, and 7-R11MH02045 to Temple University, (Philadelphia, Pennsylvania, 1966.)

[9] H. Taba, S. Levine, and F. F. Elzey, *Thinking in Elementary School Children*, Coop Res. Project No. 1574, O.E. U.S. Department HEW (San Francisco, California: San Francisco State College, 1964.)

[10] N. A. Flanders, "The Changing Base of Performance-Based Teaching," *(Phi Delta Kappa*, vol. 55, 1974), 312-315.

BEST COPY AVAILABLE

[1] R. S. Soar, R. M. Hoar, and M. Kaplonis, "Florida Climate and Control System: Observers Manual" (Gainesville: Florida University of Florida, 1971).

[2] R. S. Soar and R. M. Soar, An Empirical Analysis of Selected Follow Through Programs: An Example of a Process Approach to Evaluation, 229-258.

[3] R. S. Soar, An Integrative Approach to Classroom Learning, 1966.

[4] D. Flower, "Relation of Teacher Personality Characteristics and Attitudes to Teacher-Pupil Rapport and Emotional Climate in the E...Harbery Classroom," (Ph.D. dissertation, University of South Carolina, 1970).

[5] B. D. Brown, The Experimental Mind in Education, (New York: Harper and Row, 1968).

[6] B. S. Bloom, et al., ed., Taxonomy of Educational Objectives Handbook I, (New York: McKay, 1956).

[7] N. M. Sanders, Classroom Questions: What Kind? (New York: Harper and Row, 1966).

[8] D. D. Brown, R. Lotan, R. S. Soar, and J. N. Walsh, The Florida Taxonomy of Cognitive Behavior, (Gainesville, Florida: Institute for Development of Human Resources, University of Florida, 1972).

[9] D. S. Soar, M. Kaplonis and R. M. Soar, Florida Taxonomy of Cognitive Behavior (A-I version) (Gainesville: Florida University of Florida, undated).

[10] N. A. Flanders, Teacher Influence, Pupil Attitudes, and Achievement, Coop. Res. Monograph No. 12, (O.E. 25040, (Washington, D.C.: U.S. Department of HEW, 1965); N.A. Flanders, Analyzing Teacher Behavior, (Reading, Mass: Addison-Wesley, 1970).

[11] N. A. Flanders, Teacher Influence, Pupil Attitudes, and Achievement, Coop. Res. Monograph No. 12, (OE-25040, (Washington, D.C.: U.S. Department of HEW, 1965).

[12] A. L. Oliver, S. S. Wood, and A. Roberts, An Instructional Manual for the Reciprocal Category System, (based on a paper read at the annual meeting of the American Educational Research Association, (Chicago, Illinois, February, 9-12, 1966).

[13] D. M. Medley, The Observation, Schedule and Record, Kit Coders Manual, (Charlottesville: Virginia University of Virginia, 1974).

[14] D. M. Medley, T. J. Quick, C. G. Schutz, N. P. Anez, M. T. Carrico, and D. G. Wassey, The Personal Record of School Experience, (Charlottesville, Virginia: University of Virginia, 1973).

[15] J. A. Stallings, Follow Through Program Classroom Observation Evaluation, 1971-1972, Stanford Research Institute, Contract #OEC-0-8522460-4366(100), (Washington, D.C.: Bureau of Elementary and Secondary Education, Office of Education, Department of HEW.)

[16] R. M. W. Travers, ed. Second Handbook of Research on Teaching, (Chicago, Illinois: Rand McNally, 1973); "The Use of Direct Observation to Study Teaching," by B. Rosenshine and N. Furst, 122-183.

[17] A. Simon and E. G. Boyer, eds., Mirrors of Behavior: An Anthology of Classroom Observation Instruments, (Philadelphia, Pennsylvania: Research for Better Schools, Vols. 1-6, 1967, Vols. 7-14 and Summary, 1970).

BEST COPY AVAILABLE

8 Inservice • Needs Assessment • Competency Based Teacher Education

Gilbert F. Shearron

> Public schooling probably is the only large-scale enterprise in this country that does not provide for systematic updating of the skills and abilities of its employees and for payment of the costs involved. Teachers, we presume, can acquire marketably different ones through some process of osmosis.[1]

Most persons and organizations who give serious thought to inservice education (staff development seems to be the current term) would probably agree with the above quote from John Goodlad. One of the difficulties in inservice education has been that it only provides information to teachers — information about what someone else (usually a consultant or college professor or textbook) thought or information shared between teachers. Little or no consideration has been given to assisting teachers in developing and practicing new teaching skills in the classroom. And little attention has been given to assessing the effectiveness of what the teacher does, and what the pupils learn.

Bob Shearron has summarized inservice education by saying that often it has been long on service and short on education. It has been too fragmented, without integrated activities based upon assessed priority needs, insufficiently supported by budget, and too insignificant to leave a marked and continuing impact upon teachers and programs.[2]

Inservice education is a term that has various interpretations. For purposes of this paper, it is defined as the continuous training that a teacher needs as a result of being assigned to certain teaching situations. School districts have unique needs. Sometimes schools within school districts and classrooms within these schools also have unique needs. Few, if any,

preservice teacher education programs make the claim that their graduates have all of the teaching skills necessary to function in any circumstance. Therefore, inservice education should focus first on developing those additional teaching skills needed to implement the regular program prescribed by the school district. Secondly, the inservice program should focus on providing knowledge and teaching skills that result from change in both curricula and pupil population.

For example, in the early 1960's, teachers who were asked to teach "new math" might have needed knowledge and new methodological skills. As another example, if a school district adopts a diagnostic, prescriptive teaching strategy, then the teacher may need knowledge of diagnostic techniques as well as practice in the application of diagnostic procedures. Thirdly, inservice education aims at the need for the continuous improvement of educational personnel.

Inservice education cannot be effective without considering what needs teachers have. The position advanced in this paper is that a viable inservice education program begins with a needs assessment procedure that requires the following:

1. The identification of teaching skills needed in a particular teaching situation.
2. The development of a procedure for determining teacher needs in relation to the identified teaching skills.
3. A procedure for assisting teachers to acquire and improve his or her teaching skills.

This paper speaks to these three points. It discusses and outlines a strategy based on Competency Based Teacher Education (CBTE).

This paper does not address itself to the knowledge teachers might need. It focuses on improving teaching skills. The paper also does not deal with the question of who is responsible for teacher growth. While this is a most crucial issue, it is not within the scope of this discussion. The paper is limited to a needs assessment strategy only for classroom teachers.

Why CBTE?

Competency Based Teacher Education is a data-based system for training teachers. CBTE precisely specifies that teachers must exhibit those competencies assumed to promote pupil learning, and/or demonstrate their ability to promote desirable pupil learning.[1] Instruction in CBTE focuses directly on the development of those competencies thought to promote pupil learning. It is the opinion of the author that CBTE is a viable approach to needs assessment because it is a data-based system. Miriam Drummond points out that in a CBTE program, data are collected and used to make adjustments and changes in what and how things are done. Data are not collected for punitive reasons.[2] It is crucial that needs assessment be thought of as a procedure for improvement and not to "get teachers."

The basic elements in CBTE provide a way to develop a needs assessment system that addresses itself to the three requirements for such a system mentioned earlier in this paper. Stanley Elam identifies the essential elements of CBTE as:

1. Teaching competencies to be demonstrated are role-relevant, specified in behavioral terms, and made public.
2. Assessment criteria are competency-based, specific, mutually known, and are made public.
3. Assessment requires performance as prime evidence; takes knowledge into account.
4. Student's progress rate depends upon demonstrated competency.
5. Instructional program facilitates development and evaluation of specific competencies.

The three essential characteristics are perhaps more often thought of in the preservice realm. They can however, be applied to the needs assessment of inservice also. The position of this paper is that competency (skills based) is more easily adapted to inservice education than to preservice education.

Even if the essentials (1 - 3) carry into a needs assessment scheme, items one and two offer the opportunity to identify teaching skills in particular situations. Items three and four provide direction for establishing a procedure to determine needs, and the final item addresses itself to the improvement of teaching skills.

The systematic data-based approach utilized in CBTE provides a means to tie needs assessment and inservice together with other parts of the school program. Robert Houston and Robert Hukaspie point out that CBTE is planning in systemic terms, dealing simultaneously with the elements that comprise a total system.[1] A systematic data-based approach can make possible the continuous updating of teaching skills. The use of a systems approach however, requires that specific decisions be made as to what is to be accomplished. It also requires that problems in inservice teacher education can neither be developed nor exist in isolation from the total program of the school district. As an example of this, one might consider that if the school district is planning to provide released time or extra compensation for a teacher to improve a skill that the district itself says he or she must have, then there is little likelihood that the new skill will be developed.

The Politics of Assessment

The assessment of the needs of school personnel can be threatening to those who are being assessed. The threat may or may not be imposed. A decision to identify teaching skills presumed to be necessary for learning and the programs can become a matter of great concern to some individuals charged with carrying out the program. Organizations representing the teachers become concerned that the assessment will be used as a device to dismiss and retain teachers, rather than as a device to improve teaching.

In addition to concern about assessment, the introduction of CBTE at the inservice level can become a highly political issue. Unfortunately, CBTE often elicits an emotional response rather than reasoned consideration. Therefore it follows, between that teachers will demand a significant role in any competency based system that includes inservice education. A few of the political questions to be dealt with are: Who will determine what teaching skills are required? Who will determine which teachers need which skills? Who will determine what learning activities are appropriate to

BEST COPY AVAILABLE

A needs assessment system also should be related to the administrative hierarchy of the school district. There should be some type of subordinate-superordinate relationship. If the teacher is in the subordinate role, then the role of superordinate might be played by a principal, supervisor, or in a differentiated staffing pattern by a team leader. The superordinate, while perhaps posing a threat in some instances, does have some responsibility for seeing that programs are carried out. If needs assessment is to be based on improvement, then the superordinate has responsibility for this, being in a position to make the inservice effort more responsive to the teachers' needs. It is the position of this paper that through the use of indicators of proficiency, with the teacher and superordinate negotiating the indicators, there is less possibility of threat to the teacher. The possibility of dialogue for improvement and understanding is also expanded.

As an illustration, suppose that a competency for a teacher is in the area of diagnosis. Teacher A has no particular skills in diagnosing pupils. Teacher A and the principal agree that the teacher should focus first on diagnosing the intellectual abilities of pupils. They agree that one of the criteria to be used in assessing intellectual abilities is the use of teacher-made diagnostic tests for groups and individuals. In order to become more proficient, Teacher A is to develop a skill in building tests that can identify where pupils are on specific skill continuums in reading and mathematics. Inservice activities should then be recommended by the principal that will help Teacher A to learn how to construct and administer diagnostic tests in reading and mathematics. Once the teacher develops the skill and utilizes it in the classroom, then there would be opportunity to develop further diagnostic skills that would continue to make the teacher more proficient.

We believe that teachers and administrators can work together in identifying needs if it is for purposes of improvement. Obviously, there are the human requirements of trust and consideration that must be present in the relationship for this or any other working relationship to succeed. We are hypothesizing that if a subordinate and a superordinate agree on an area of improvement, then precisely define what the improvement will be, there is a good possibility that improvement can be assessed by both parties. Again, it is also important to consider that the administrative hierarchy be responsive to the needs of the teacher. A hierarchy is likely to react more readily if it is represented in the process.

Summary

This paper has presented a proposed strategy for developing and utilizing a needs assessment strategy based on Competency Based Teacher Education. The strategy takes into account both conceptual and political considerations. It comes from attempts to combine some of the conceptual elements of CBTE with pilot efforts to use these conceptual notions in school districts. Some of the suggested procedures are as follows:

1. Teachers should make the final decisions on needed competencies to perform effectively in a school district. However, teachers should

BEST COPY AVAILABLE

have opportunities to consider the theoretical aspects of teaching, including research findings prior to making decisions about needed competencies.

2. The assessment procedure should begin with the assumption that needs assessment is for the improvement of teaching. Therefore, we begin with the notion that teachers are competent, and that the purpose of the needs assessment is to help them become more proficient.
3. A continuum of proficiency for each competency should be developed in order to allow for continuous teacher growth. The indicators of proficiency on this continuum should be negotiated by the assessor and the teacher.
4. Self assessment should be part of the needs assessment procedure. However, self-assessment skills need to be developed in order to make this a successful venture.
5. The administrative hierarchy of the school district should be involved in the assessment procedure.

FOOTNOTES

[1] John I. Goodlad, "The Schools vs. Education," *Saturday Review*, (April 19, 1969), 59-61.

[2] Bob Bhaerman, *A Paradigm for Accountability*, American Federation of Teachers Quest Paper, (No. 12, August, 1970), 3.

[3] Gilbert F. Shearron and Charles E. Johnson, "A CBTE Program in Action: University of Georgia," *Journal of Teacher Education*, 24 (Fall, 1973), 187.

[4] William H. Drummond, "Comments on Achieving the Potential of PBTE," Appendix B-3, *Achieving the Potential of Performance-Based Teacher Education: Recommendations*, (Washington D.C.: American Association of Colleges for Teacher Education, December, 1971), 39.

[5] Stanley Elam, *Performance-Based Teacher Education, What is the State of the Art?* (Washington, D.C.: American Association of Colleges for Teacher Education, December, 1971), 18.

[6] W. Robert Houston and Robert B. Howsam, *Competency-Based Teacher Education: Progress, Problems, and Prospects*, (Chicago: Science Research Associates, Inc., 1972), 11.

[7] Theodore E. Andrews, *Atlanta or Atlantis?* A publication of the Multi-State Consortium on Performance Based Teacher Education.

[8] "Atlanta's Elementary Curriculum Revision Project," *Your Schools*, a publication of the Atlanta Public Schools, (April, 1973), 3.

[9] Ibid.

[10] *An Assessment System Utilizing Teacher Competencies for Differentiated Staffing*, Developed by Mentor Ohio Exempted Village School District and The Center for Competency Based Education, University of Georgia, (Athens, Georgia: February 1974).

[11] *Model Teacher Education and Differentiated Staff Assessment*, Mentor Ohio Exempted Village School District, ESEA Title III Project (45-71-208-2).

[12] Barak Rosenshine and Norma Furst, "Research on Teacher Performance Criteria," Chapter 3, in B. O. Smith (ed.) *Research in Teacher Education*, (Englewood Cliffs, N. J.: Prentice Hall, Inc., 1971).

[13] W. Robert Houston, *Strategies and Resources for Developing a Competency Based Teacher Education Program*, (New York State Education Department and Multi-State Consortium on Performance-Based Teacher Education), 21.

[14] *Ibid.*, 21-22.

[15] John D. McNeil and W. James Popham, "The Assessment of Teacher Competence," *Second Handbook of Research on Teaching*, (Chicago: Rand McNally College Publishing Company, 1973), 231-2.

BEST COPY AVAILABLE

9 Assessing Teacher Competence

Beatrice A. Ward
Greta Morine
David C. Berliner

Assessment of teacher performance may be conducted for a variety of reasons, such as selection for entrance into preservice training, identification of the types of training most needed by a preservice and/or inservice teacher, awarding of a certificate to teach, and research into effects of teacher use of particular skills.

For the purposes of this paper, we are focusing upon assessment as a means for selecting training options. We have made this choice for two reasons. First, the body of knowledge in teacher effectiveness research at the present time is too limited and too subject to question to justify using assessment data for selecting preservice candidates or granting licenses to teach.[1] We do not as yet know enough about which aspects of teaching are essential for affecting student performance positively to use specific levels of teacher performance as "go" or "no go" criteria for entry into or promotion within the profession.

Yet much is known about measurement of teacher performance. Large numbers of evaluative studies report data describing typical levels of teacher performance before and after training.[2] To ignore these data while waiting for completion of effectiveness research is, in our opinion, as great a mistake as using them for hiring or firing purposes. Thus, our second reason for concentrating upon assessment for training purposes is to illustrate how the current state-of-the-art in the study of teaching can be productively applied to the design of a self-improvement program for an individual teacher or a group of teachers.

Within the spectrum of performance based teacher education, several training approaches have been developed. The assessment techniques described herein illustrate two diverse points on the continuum of ap-

BEST COPY AVAILABLE

proaches. The skill diagnosis approach represents a behavioristic view of assessment, the inquiry approach, a humanistic view. Three critical differences in the approaches are: (a) the data base from which decisions are made regarding the need for training, (b) the individuals responsible for designing the assessment event, and (c) the type of evaluative feedback provided.

Both approaches are under development at the Far West Laboratory for Educational Research and Development. Work on the performance tests of teaching is in cooperation with National Teacher Corps and Corps sites at Kansas State Teachers College, Emporia, Kansas; Buffalo State University College, Buffalo, New York; and the Southern Consortium of Colleges and Universities, Mobile, Alabama. The inquiry approach is being developed in cooperation with the Oakland Unified School District, Oakland, California, under the auspices of the Bay Area Teacher Training Complex, sponsored by a grant from the U. S. Office of Education.

A Skill Diagnosis Approach

The skill diagnosis approach to analysis of competency in teaching competency is based on a belief that teacher trainers and researchers need to introduce a medical model into the way they examine teachers and their teaching practices. Within the medical profession, the model requires the monitoring of a number of potentially important systemic indicators, such as heart rate, patellar reflex, and blood pressure. Then clinical knowledge about the client is developed—his or her job requirements, temperament, previous history, etc. On the basis of these data, if certain indicators are out of tolerance, individualized prescriptions are made (e.g., lose weight, quit smoking, take medication four times a day).

Teacher trainers and researchers also need to develop ways to monitor the systemic functioning of their clients, perhaps checking their warmth instead of their temperature and their response to pupils rather than to light. Then, with clinical insight derived from information about the teacher's experience and interests—and perhaps the age and social class of the students to be taught—individualized prescriptions could be prepared, such as: "Take Module 014 from the Houston Module Bank for learning reinforcement techniques to use with young inner-city children."

How to find the indicators for diagnosis and prescription of this kind is the concern of the skill diagnosis approach to assessment.

Knowing "What" and Knowing "How"

Multiple choice or essay tests are easy to develop and grade, can sometimes even be reliable, and cost very little per client tested. One can readily assess knowledge about *what* discovery-teaching is, *what* one should do with obscene notes, *what* the definition of contingent reinforcement is, etc. This knowledge of "what" may be important, but its relationship to knowing *how* to teach is quite indeterminate. A recent view correlating a highly reliable test of knowledge about teaching with various measures of teaching performance informs us that virtually no relationship exists between the two variables.[1] Indicators for examining teaching should not come from

BEST COPY AVAILABLE

paper-and-pencil tests which measure what is known. The indicators must be based on how teachers perform.

Observation in Natural Environments

Some believe, along with the many developers of observation instruments, that observations of how teachers perform in their classrooms is the way to assess teaching competency. But initial observations might be very much like the annual visit to the classroom by the principal—an exercise in futility. In most philosophic systems, the "good" teacher is one who varies his or her behavior depending upon the child, the learning objective, the time of year, and a host of other variables. All these are influences on teacher behavior designed to make the behavior unstable in situ. Thus observation of teachers in natural environments may be an inappropriate way to develop reliable and valid indicators of teaching performance.

Performance Tests

In performance tests of teaching, where a controlled environment exists, indicators for skill diagnosis may be found that are similar in type to those used by our medical colleagues. In these diagnostic tests, a teacher trainee receives a statement of two or three objectives, some written material on a topic, and sample test items. The teacher is told to prepare to teach five or six students for 20 minutes. The teaching performance is video- or audio-taped. Frequency counts or ratings on as many as 50 variables are compiled.

In order to develop such tests, three important elements must be considered.

Choosing variables. Thus far, the literature on teacher effectiveness (such as it is) along with the value systems of the participating researchers and teacher trainers have directed the selection of variables. Variables also have been picked to match those included in existing teacher training materials since it is important to examine the pharmacy, i.e., the Minicourses, texts, protocols, simulations, etc., available for use in responding to a skill diagnosis before writing a prescription. In the performance tests, we try not to be like the doctor who diagnoses terminal cancer and offers no cure. We assess first if there is a cure, i.e., a training module, and then proceed to diagnose by examining some indicator to see what, if any, training need exists in the curable area.

Evaluating performance. The teacher's performance is evaluated against (a) norms which have been accumulated from other trainees, (b) the performance of master teachers attempting the same lesson, and (c) criteria we have set for performance in a particular area. In light of each or all of these sources of information about teaching performance, a diagnosis is made. Feedback is provided to the trainees on each diagnosis. For example:

Variable 1: Teacher gives specific praise or encouragement.

Variable Description: The teacher praises or encourages students and clearly states why praise or encouragement is given. Example—"That is

a good answer because

Performance Evaluation

1.) This evaluation occurred 3 times in your teaching.

2.) Based on norms and our criteria, 3 to 12 times is desirable.

3.) The diagnosis to be made is

_____ Performance is within limits

_____ Performance is too high

3 Performance is too low

Prescriptions: If performance is too high, see Doc. 2, 4, or 21
If performance is too low, see Rec. 20, 1, 32

This form of feedback is continued for as many variables as desired. In each case, the trainee receives descriptive information and definitions of the variable, his actual performance, estimates of desirable performance, and prescriptions for remediation.

Prescription. The numbers in the prescription list refer to training materials which exist in the 'pharmacy.' For example, Rx: 20.1 may be the code for a lesson in a Minicourse that focuses on contingent reinforcement. Rx: 32 may refer to a protocol videotape that demonstrates (i.e. concept of "because extensions" in communication, or the trainee may be referred to a chapter in a text on the effects of praise and blame on performance. If the trainee's performance was too high, that is, if performance exceeded the level considered desirable, the trainee might be directed to materials which teach about habituation, since praise will become ineffective if too much is given, because people tend to disregard it or question its genuineness.

For any teacher variable of interest, if an out-of-tolerance situation is detected in a performance test, one or more prescriptions can be recommended. In this way these tests are as diagnostic as the physician's examination. An individually prescribed program of training may be developed which allows a teacher to concentrate on areas where training will be most useful and skip areas where performance competency has already been demonstrated. Preliminary study of these performance tests has shown them to be reliable. Stability of performance seems to hold over parallel lessons and for the same lesson taught at different times. This means that the performance tests, when thoroughly developed, can be used not only to plan training but also as objective measures of the effectiveness of competency-based teacher training programs.

An Inquiry Approach

All approaches to increasing teacher competence share at least one common purpose—the improvement of instruction. The two approaches discussed here share a common orientation to attainment of that goal—the view that formative evaluation of teacher effectiveness is the most viable route. The inquiry approach differs from the diagnostic-prescriptive approach in actively pursuing the following additional purposes:

1. To provide a supportive climate in order to maximize change in teacher behavior.
2. To individualize the inquiry project criteria on which assessments of teacher competence are based.
3. To provide teachers with a strategy for continuous self-assessment and self-improvement.

The inquiry approach to be described here has as its focus a re-analysis of the process of relating into one's own teaching behavior proposed by Flanders.[1] This analysis was initially related to assessment of teacher competence in a data paper on the California Beginning Teacher Evaluation Study[2] which used the Flanders system as one method of assessment. The comparative procedure which is an essential element of this approach was tested and found to be effective in getting teachers to be more specific in identifying training needs and more interested in continued self-training.[3]

Procedure for Inquiry Projects

A teacher embarking upon an inquiry project for the purpose of self-assessment would engage in the following steps:

1. Identify one or more instructional objectives and plan two alternative lessons to achieve the objectives.

   Examples of plans that carry some one instructional objective may be simple and unsophisticated, phrased in such as "to identify correctly the words on inquiry graph in list of ten nouns and verbs." The objective may be more complex and long-term, such as "to increase pupil skill in listening to and reacting to each other's comments." The alternative lessons for achieving the instructional objective may involve simple variation in use of one particular teaching skill, such as phrasing pupils for every correct answer in one lesson and praising at intervals for attentive behavior in the other lesson. Variation in instructional strategies would be a more complex approach in planning alternative lessons, e.g., using an inductive discovery strategy in one lesson and a deductive strategy in the other. (The illustrative inquiry project in the appendix has short-term objectives and involves variation in the use of a particular skill as well as variation in instructional strategies.)
2. Decide what data to collect for evaluation of the effectiveness of the two lessons.

   Data on the interaction behavior of the teacher will usually be used to determine whether the planned variations in the two lessons did in fact result in observable differences. Data on the interactive behavior of pupils may be used to determine if instructional objectives were achieved. Samples of pupil work or pupil attitude inventories may also be used for this purpose. Interaction data may be collected by use of a standard category system, but usually the teacher will need to adapt such a system to highlight the differences he/she plans to create. It is also possible for the teacher to design his/her own simple category system. (The sample inquiry project in the appendix utilizes the Flanders Interaction Analysis System and

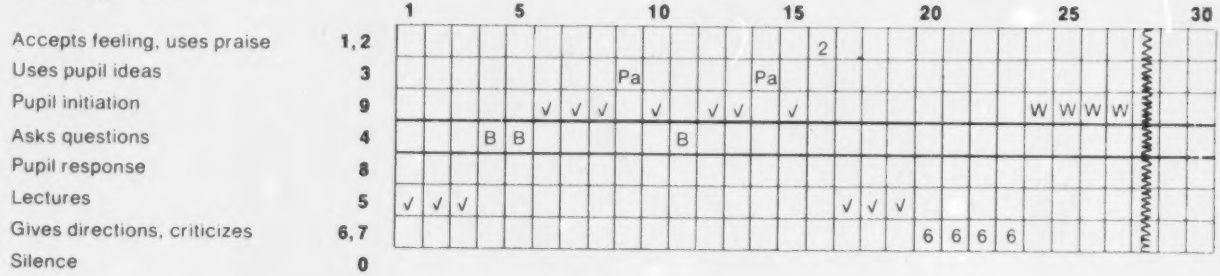
[redacted]

[redacted]

Lesson One (concluded) ". . . Historian"

Planned Procedure: 3. Ask pupils to predict changes that might occur in families in 2020 if the trends (similarities and differences) noted from 1890 to 1970 continue. Paraphrase their answers to clarify ideas expressed. Make writing assignment.

Coded Segment III



Special Subscripts: 3-Pa- Teacher paraphrases pupil answer to clarify.

9-W- Pupils write answers at seats.

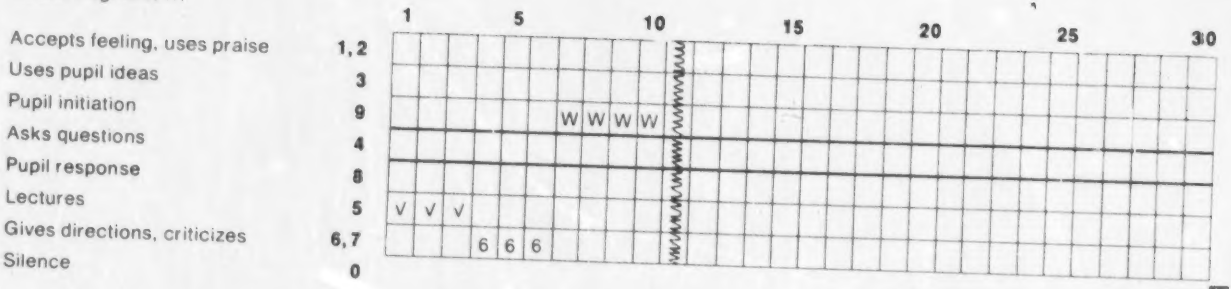
2 - Teacher uses praise.

BEST COPY AVAILABLE

Lesson Two (concluded) ". . . Science Fiction Writer"

Planned Procedure: 3. Make writing assignment.

Coded Segment III



Special Subscripts: 9-W- Pupils write answers at seats.

BEST COPY AVAILABLE

BEST COPY AVAILABLE

10 Teacher Competency Research

Barak Rosenshine

The purposes of this paper are to review current research on teaching skills and to suggest procedures for future research which can take place as part of competency based teacher education.

There are two action suggestions within this paper. The first is the suggestion that each institution engaged in competency based teacher education conducts one experimental study per year involving teacher skills and student achievement. The second suggestion is that each institution produces one interpretative review on a teaching skill per year. Details about how this might be done and possible topics for study are contained within this paper.

State of Research on Teaching Competencies and Pupil Achievement

Most papers on teacher education contain the embarrassing recognition that the present scientific base for teaching and teacher education is primitive. That is, the number of studies which have looked at both teacher behavior and student outcomes is embarrassingly small. A diligent search will uncover less than a hundred studies. The quality of many of these studies is questionable. The results are most useful for suggesting future research; the results of these studies are *not* sufficiently strong or clear to direct teacher training practices or certification or evaluation of teachers.

In contrast, there is no lack of research on *how to train* teachers to introduce a lesson, ask questions at higher cognitive levels, redirect pupil responses, praise students, or encourage student talk. Nor do we lack for

BEST COPY AVAILABLE

lists of teaching skills. What is lacking, however, is research demonstrating an experimental or correlational connection between these skills and measures of change in pupil achievement and affect.

Statements on the lack of research may be surprising to many readers. Readers are invited to check these claims for themselves. Consult a textbook in educational psychology, teaching, or subject area methods and see how much research exists for oft-proclaimed teaching procedures such as matching teacher style and pupil style, asking questions, using behavioral objectives, promoting student talk, and motivating students. In these areas, as in others, I find only a few studies which took place in classrooms, and these have yielded equivocal results.

I believe that there are insufficient studies to date, to permit an evaluation of the promise of this research. Currently, an important task is to generate more and improved research so that the promise of this research and directions for future study can be evaluated more adequately.

Research Findings to Date

There have been very few attempts to summarize the current state of our knowledge on teaching competencies. In one review, the correlational research on some 25 teaching competencies is presented.[1] Nine of these 25 variables were selected as the most promising for future research on the basis of results obtained to date. A subsequent review supports, in general, the Rosenshine-Furst findings and conclusions.[2] It should be noted that these nine "most promising" variables were selected from a review of the correlational literature, and future experimental studies may not support all of these correlational results.

The variables below were listed in the 1973 summary, and more details are available in the 1971 volume. Those nine variables were:

1. Clarity of teacher's presentation
2. Variety of teacher-initiated activities
3. Enthusiasm of teacher
4. Teacher emphasis on learning and achievement
5. Avoidance of extreme criticism
6. Positive responses to students
7. Student opportunity to learn criterion material
8. Use of structuring comments by teacher
9. Use of multiple levels of questions or cognitive discourse

These nine variables are promising suggestions for future research. They cannot be used as checkpoints to assess teacher competency because the research to date is too incomplete.

At first glance, the above list of most promising findings may appear to represent educational platitudes. Their value can be appreciated, however, only when they are compared to the variables which have not shown signifi-

BEST COPY AVAILABLE

Lack of consistent relationships with achievement to date. These variables are

1. Non-verbal approval
2. Praise
3. Teacher warmth
4. Ratio of all indirect to all direct behaviors (I/D ratio)
5. Number of different teacher behaviors (flexibility)
6. Number of questions
7. Questions classified into two levels
8. Teacher talk
9. Student talk
10. Teacher experience
11. Teacher attitude toward subject matter

Against these above two lists, there is a much larger list of teacher competencies for which training is currently given. For most of these variables I have been unable to locate more than one or two classroom-based studies.

These variables include:

1. Probing of student responses
2. Selection of tasks to fit readiness level of pupils
3. Motivating pupils' attending behaviors
4. Establishing and maintaining rapport
5. Providing feedback
6. Conducting individual activities
7. Presenting information
8. Explaining subject matter
9. Giving examples
10. Using inductive strategies
11. Using deductive strategies
12. Pacing
13. Responding to student answers
14. Preventing discipline problems
15. Giving homework
16. Solving problems
17. Developing pupil self-concept
18. Developing learning to learn skills
19. Developing acceptance of responsibility
20. Selecting instructional strategies
21. Involving students in self-evaluation

The state of the art, then, is one of a lack of research. We simply do not have enough studies about the relationship between teaching competencies and pupil achievement.

BEST COPY AVAILABLE

Teacher Questions

The situation is more complex than merely a lack of research. There is also a need to integ-ate research findings into some meaning for teacher educators. At times this is a difficult task because of the eclectic of studies and the difficulty of comparing studies. These difficulties are illustrated in the mini-review on teacher questions presented below.

Training preservice teachers to ask a variety of types of questions is part of any teacher training program. In some programs, instruction on types of questions is given through assigned readings or through worksheets and examples. In other programs, trainees receive additional practice in a microteaching setting and receive feedback on the number of questions they asked which were factual, analytic, divergent, convergent, comprehension, synthesis, evaluative, application, comparative, inferential, or any of the bewildering variety of categories into which questions have been classified. The objective of the training, of course, is to increase the teacher's use of a variety of nonfactual questions. Tens of training studies have been conducted, and almost all show that a teacher's use of a variety of "higher cognitive level" questions increases as a result of training.

Although the results of training studies are straightforward, the results are not so clear when one inspects correlational and experimental studies involving teacher questions and student achievement. (To the best of my knowledge, interpretative reviews of teaching behaviors and other student outcomes such as student attitudes or beliefs about oneself have not been written. Such additional reviews are urgently needed, and readers are invited to consider conducting such reviews.)

Correlational Studies on Questioning

Some 15 correlational studies were found which contained classifications of teacher questions and measures of student gain. Fifteen studies is an unsatisfactorily small number when one considers that there are at least five grade level groupings (i.e. primary, intermediate, junior high, senior high, and college), at least four subject areas, and at least two types of cognitive outcome measures (factual and comprehension). This alone yields 40 separate cells -- and there are only fifteen studies, and the quality of many of them is less than desired.

Fifteen correlational studies on teacher questions and student gain reported in the last 15 years of research? Without any trouble, one could easily find 30 books, chapters in books, and journal articles, which claim to teach how to ask questions. In the two volumes on Resources for Performance-based Education there are 36 modules on how to ask questions -- and an equal number of filmstrips, movies, audiotapes, and multimedia kits. (We are instruction-rich and data-poor.

Differences in definitions. Even with only 15 studies, there are problems in interpretation because variables with the same title (e.g., factual questions) are defined and coded differently in different studies. Although investigators usually classify questions into two types, many questions

BEST COPY AVAILABLE

which are classified as higher level questions. (A table of intercorrelations between the eight types of questions can be obtained by writing to the author.) Questions which require a student to explain how he got an answer or questions which require the student to a thought problem have been classified as factual by this investigator and as higher level questions by another.

Each of the three investigations which investigated used to classify questions are not known how many beginning questions, for example, were classified among the higher level questions by one investigator and among the lower level by another. Because of these problems, any comments on the level of results must be treated with caution.

Results of correlational studies. Of the six studies in primary grades in which questions were broken into two types, significant results favoring use of higher level questions were not obtained in any study. These results were for reading achievement, science achievement tests, and author-developed tests. In one study, open questions and a nonacademic focus were significantly negatively related to pupil gain in the first grade but not in the third grade for control class participants. In another study, questions which asked children to give a personal preference or response from personal experience were negatively related to gain in reading participants for middle class children. In the same study, these results are not true for mathematics gain. Thus, the overall results for open questions in the primary grades are either nonsignificant or negative. In four of the eleven studies there was no significant relationship between the frequency of factual questions and student achievement. In the fifth study there was a positive significant relationship. In the sixth study, a mixed and positive and significant for academic factual questions the correlations were negative and significant for nonacademic factual questions. Thus, the overall results for factual questions are nonsignificant but with a mixed (some positive relationships)

with mixed toward positive relationships.

Studies. The four studies in the intermediate grades are more difficult to interpret than those in the primary grades because of greater differences in coding and statistical procedures. As in the primary grades, there is a trend for factual questions to be positively related to achievement gain and for open questions which concern student personal experiences to be negatively related. In addition, one study showed a significant curvilinear relationship between instructional questions and student achievement, but these results were not replicated in the other three studies. Indeed, in another study, there was a moderate number of divergent to convergent questions which were significant. Of the remaining two studies, one yielded all non-significant results and one found a negative trend between questions regarding student interests and opinions and student achievement. This last finding appears similar to that observed in two of the studies conducted in the primary grades.

In two studies in junior high school grades, both in science, each used a ratio of higher level to lower level questions. In one, a moderate ratio was most functional, and this finding appears similar to that observed in intermediate grade complexity. However, both sets of results must be taken as

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]

[illegible]

[illegible]	[illegible]
1. [illegible]	[illegible]
2. [illegible]	[illegible]

[illegible]

BEST COPY AVAILABLE

of Education.

I would propose that in addition to training teachers in the lists of competencies developed by a committee, we *also* help teachers study these different programs and then train them in the competencies necessary to implement these different types of programs. In other words, why not bring curriculum and instruction together, instead of leaving curriculum as a junior partner to teacher training in teaching skills.

Curriculum Materials Packages Within Preservice Training

How might curriculum packages be used in preservice education? In elementary mathematics education, for example, we don't know which curriculum product, if any, the teachers will be using throughout their careers. However, as illustrated above, one can group the existing packages and the ones being developed into two, three, or four types and train teachers in the skills necessary to use these *types* of materials. Preservice teacher training can be conceived of as training teachers in the skills required by the *types* of educational products that are expected to be in use during the early part of their careers. Preservice teachers would learn the skills by using and studying the different types of curriculum packages. In elementary mathematics, for example, teachers would learn skills of informal teaching by attempting to implement the mathematics portion of the Bank Street or the EDC models; they would learn skills of structured mathematics teaching by learning to implement programs such as the Englemann-Becker or the Karnes, and they would learn skills of "modern math" by working through sections of the Greater Cleveland Math Programs or the Madison Math Program.

Curriculum Programs and Accountability

If curriculum materials packages are being used (and if development of new materials and programs is assigned to appropriate agencies), then the responsibilities are reallocated among the various groups. Teachers are responsible for implementing the program according to the instructions of the developers. If teachers decide to change the instructions, then those teachers become accountable for the success or failure of the students to learn. The developers are responsible for determining the most appropriate instructional procedures for the students whom the teachers will have, and if the teacher follows the instructions, then the developers are responsible for the outcomes. The developers and the teacher educators are responsible for providing the teacher training; the administrators are responsible for making the training available and for providing sufficient teacher training; and the public is responsible for providing the funds for the materials and the training. If the program includes parents working with the students, then the parents are responsible for their contribution.

In other words, rather than each group trying to fix the blame on the other, in this situation each group is responsible for facilitating the work of the other.

BEST COPY AVAILABLE

This proposed realignment of responsibilities seems analogous to that of a physician who is responsible for using the best machinery and available knowledge in his diagnosis and treatment, but is not responsible for curing heart disease and cancer. The responsibilities for finding cures and drugs belong to special and appropriate research organizations. Of course we want people cured of cancer; but will it be done by setting down a list of "performance competencies" for physicians? Of course we want more human interaction; but will it be done by setting down a list of "performance competencies" for teachers?

Needed Research on Curriculum Packages and Teaching Competencies

The above remarks on the usefulness of curriculum packages need to be tested. One needs to research whether the teaching skills which are part of the curriculum packages are indeed related to student achievement. At present, there have been only a few studies in this area. In one set of correlational and experimental studies, significant results were obtained on the importance of teacher correction procedures in a highly structured curriculum program.[32] In other studies using other variables, the results have not been clearcut.[33]

Summary

The major purpose of this paper is to call for more research on teacher competencies and more reviews on the results of previous studies. Because of the importance of this research, institutions involved in teacher competency training are urged to conduct one study and one review of research per year. If such work is initiated, then within two years we will have a clearer knowledge base from which to proceed in training and research. Suggestions for conducting such research are provided within the paper.

FOOTNOTES

[1] B. Rosenshine, *Teaching Behaviors and Student Achievement*, (New York: Humanities Press, 1971).

Summarized in B. Rosenshine and N. Furst, "Research on Teacher Performance Criteria," B. O. Smith (Ed.) *Research in Teacher Education*, (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1971).

Updated in B. Rosenshine and N. Furst, "The Use of Direct Observation to Study Teaching," Chapter 5 in R. M. W. Travers, *Second Handbook of Research on Teaching*, (Chicago, Illinois: Rand McNally and Co., 1973).

[2] D. A. Potter and M. Lockhead-Katz, *A Critical Review of the Literature: Teacher Performance and Student Growth*, (Princeton, New Jersey: Division of Teacher Education, Educational Testing Service, 1973).

[3] W. Robert Houston, et al., *Resources for Performance Based Education* (Albany, New York: State Education Department, 1973).

W. Robert Houston, et al., *Resources for Performance Based Education Supplement A* (Albany, New York: State Education Department, 1974).

[4] G. R. Thompson and N. C. Bowers, *Fourth-Grade Achievement as Related to Creativity, Intelligence, and Teaching Styles*. Paper presented to the Annual Meeting of the AERA, 1968.

[5] H. V. Perkins, "Classroom Behavior and Underachievement, *American Educational Research Journal*, (1965) 2, 1-12.

[6] R. Spaulding, *Achievement, Creativity and Self-Concept Correlates of Teacher-Pupil*

BEST COPY AVAILABLE

Transactions in Elementary Schools. (Urbana, Illinois: The University of Illinois, U. S. Office of Education, Cooperative Research Project No. 1353).

[7] J. Stallings, et al., *Evaluation of Follow Through Models*, 1972-73. (Menlo Park, California: Stanford Research Institute, 1974).

[8] J. Brophy and C. Egbertson, *Appendix to First Year Data of Texas Teacher Education Project*. (Austin, Texas: University of Texas Research and Development Center in Teacher Education, 1974).

[9] G. R. Thompson and N. C. Bowers, *Fourth-Grade Achievement as Related to Creativity, Intelligence, and Teaching Styles*, 1968.

[10] G. Kleinman, *General Science Teachers' Questions, and Pupils' Understanding of Science*. (University Microfilms No. 65-3961, 1964).

[11] N. F. Furst, *The Multiple Languages of the Classroom*, (Philadelphia, Pennsylvania: College of Education, Temple University, 1967).

[12] D. Solomon, et al., "Teacher Behavior and Student Learning," *Journal of Educational Psychology*, 1964, Vol. 55, 23-30.

[13] Greg Millett, *Comparison of Four Teacher Training Procedures in Achieving Teacher and Pupil Translation Behaviors in Secondary School Social Studies*. Paper presented to the American Educational Research Association, 1969. (Austin, Texas: University of Texas. ED027256).

[14] G. L. Miller, "Collaborative Teaching and Pupil Thinking," *Journal of Teacher Education*, (1966): 17, 337-358.

Millett, *Comparison of Four Teacher Training Procedures in Achieving Teacher and Pupil Translation Behaviors in Secondary School Social Studies*, 1969.

V. M. Rogers and O. L. Davis, *Varying the Cognitive Level of Classroom Questions: An Analysis of Student Teachers' Questions and Pupil Achievement in Elementary Social Studies*. Paper presented to the American Educational Research Association, March, 1970. (Lexington, Kentucky: University of Kentucky).

[15] W. L. Hutchinson, "Creative and Productive Thinking in the Classroom," *Journal of Creative Behavior*, (Fall, 1967), 1, 419-427.

R. L. Davidson, "The Effects of an Interaction Analysis System on the Development of Critical Reading in Elementary School Children," Classroom Interaction Newsletter, (May, 1968) (Doctoral Dissertation, Denver, Colorado: University of Denver, 1967), 12-13.

[16] J. L. Ciesla and J. R. Okey, *The Effects of Training Preservice Teachers to Use Bloom's Mastery Teaching Strategy*. (Bloomington, Indiana: School of Education, Indiana University, 1974).

[17] Willavene Wolf, et al., *Critical Reading Ability of Elementary School Children*. U. S. Office of Education, Project No. 5-1040. (Columbus, Ohio: 1967).

[18] W. W. Lynch, C. Ames, C. Barger, W. Frazer, S. Hillman, and S. Wisehart, *Effects of Teachers' Cognitive Demand Styles on Pupil Learning*, 1973.

[19] L. J. Buggey, *A Study on the Relationship of Classroom Questions and Social Studies Achievement of Second Grade Children*, Doctoral Dissertation. (Seattle, Washington: University of Washington, 1971). (UM 71-28, 385).

[20] T. V. Savage, *A Study of the Relationship of Classroom Questions and Social Studies Achievement of Fifth-Grade Children*. (University Microfilms, UM 72-78, 661, 1972).

[21] L. J. Buggey, *A Study of the Relationship of Classroom Questions and Social Studies Achievement of Second Grade Children*, 1971.

T. V. Savage, *A Study of the Relationship of the Classroom Questions and Social Studies Achievement of Fifth-Grade Children*. (University Microfilms, UM 72-78, 66, 1972).

[22] Wolf and others, *Critical Reading Ability of Elementary School Children*, 1967.

[23] Lynch, Ames, Barger, Frazer, Hillman, and Wisehart, *Effects of Teachers' Cognitive Demand Styles on Pupil Learning*, 1973.

[24] Ciesla and Okey, *The Effects of Training Preservice Teachers to use Bloom's Mastery Teaching Strategy*, 1974.

[25] N. R. Dodl, *The Florida Catalog of Teacher Competencies*, (Chipley, Florida: P. O. Drawer 190, 1973).

W. R. Houston, et al., *Resources for Performance Based Teacher Education*, (Albany, New York: State Department of Teacher Education and Certification, 1973).

W. R. Houston, et al., *Resources for Performance Based Education Supplement A*, (Albany, New York: State Education Department, Division of Teacher Education and Certification, 1974).

BEST COPY AVAILABLE

[1] N. L. Flanders, *Analyzing Teacher Behavior*. (New York: Addison-Wesley, Inc., 1970).

[2] M. Gall, Study of Teacher Higher Cognitive Level Questions. (San Francisco, California: Far West Laboratory for Educational Research and Development, 1974).

[3] Lynch, Ames, Barger, Fraser, Holman, Wareham, *Effects of Teachers' Cognitive Demand Styles on Pupil Learning*, 1973.

Cissie and Okeu, *The Effects of Training Preservice Teachers to use Bloom's Mastery Teaching Strategy*, 1974.

[4] C. J. Wright and G. Nuthall, "Relationships Between Teacher Behaviors and Pupil Achievement in Three Elementary Science Lessons," *American Educational Research Journal*, (1970, Vol. 7), 477-493.

[5] Instructional Appraisal Services, *Teaching Improvement Kits*, (Los Angeles, California: 1972).

[6] D. G. Johnson and J. W. Bommarito, *Tests and Measurements in Child Development*, (San Francisco, California: Jossey-Bass, Inc., 1971).

[7] M. A. Siegel and S. Rosenshine, *Teacher Behavior and Student Achievement in the Bereiter-Englemen Follow Through Program*. (Urbana, Illinois: Computer Education Research Laboratory, 1974).

M. A. Siegel, *An Experimental Investigation of Teacher Behavior and Student Achievement in the DISTAR Instructional Program*. (Urbana, Illinois: Computer Education Research Laboratory, 1974).

[8] D. E. Feinlob, et al., *The Effects of Three Levels of Training on the Implementation of a Preschool Program*. Paper presented to the Annual Meeting of the American Educational Research Association, 1974.

J. Stallings, et al., *Evaluation of Follow Through Models, 1972-1973*. (Menlo Park, California: Stanford Research Institute, 1974).

BEST COPY AVAILABLE

PART III

PROGRAM EFFECTIVENESS

BEST COPY AVAILABLE

11 If the Program is Competency Based, How Come the Evaluation is Costing so Much?

Richard Jensen

Introduction

A competency based program is presumably thoroughly developed by the time it is in full-scale operation, which presumably means that it gets its graduates up to the designed level of quality and quantity of achievements. (We pass lightly over the fact that it may do this by lowering the design level in the light of realistic hard trials.) Here at last, one feels — or is it just that one hopes? — is a program that knows what it does and says what it knows. Here we have laid to rest the misleading advertising of inflated phrases and vague promises that pervade traditional education catalogs.

Apart from the moral and scientific charms of the competency based approach long extolled in prose and poetry (somehow the poetry seems to creep back into the catalogs), there does seem to be at least one obvious pecuniary advantage about this innovation, namely the elimination of any need for fancy evaluation. Competency based education, the honest fellow means this — or is it her — proclaims on its sleeve. They need no embellishment.

I wish I could add to your enjoyment of this evening's banquet by encouraging you to believe that the competency based path is indeed rewarded in this way. But I propose instead to contribute to mitigation by stressing that the direct cost and quality of evaluation in competency based programs that I know about are not even in the ballpark for adequacy. Not. I hasten to add, that this puts it in a good position compared to most educational programs. While it is true that the competency based approach offers the possibility of a modest saving in one dimension of evaluation, this saving is usually not realized or not realized fully, and competency based evaluation

BEST COPY AVAILABLE

12 Task Force on Evaluating Program Effectiveness

David Potter
Horace Aubertine

Competency based teacher education is an operational reality at relatively few institutions. At most other teacher education institutions, the competency based movement is an emergent phenomenon that is proceeding at various rates. Evaluation of such emergent competency based programs must, of course, fulfill all of the usual functions of program evaluation; it must provide objective, impartial measurement of the effects of the program. In addition, program evaluation should form the core of what is in effect a management information system designed to provide program decision makers with accurate, objective, timely information concerning the relative effectiveness of all elements of the system. Such evaluative feedback is essential for effective management; without it, crucial decisions often will be made on the basis of hunch or expediency.

Ideal program evaluation as seen by the Task Force members measures program success against the 13 criteria discussed by Scriven during this conference. These criteria rest on what is essentially a systems approach to evaluation—an approach which considers not only the extent to which the organization is successful in reaching its stated goals, but which also considers things which are in a sense subordinate goals (for example, the level of morale among people involved in the program). Program success is seen as a complex, multidimensional concept. The evaluator should study not only the degree to which the program achieves its goal of improving education by improving the preparation of educational personnel, but also the processes through which the organization functions.

At the most basic level, the keystone of performance based education

165

BEST COPY AVAILABLE

programs is the willingness to be held accountable for achieving stated program objectives through the use of training and assessment procedures which are also open for public inspection. The openness to summative evaluation implied by this may be at once performance based education's greatest strength and its greatest weakness. The strength lies in the willingness to be evaluated on the basis of publicly stated objectives; but it is also a weakness in that premature summative evaluation could result in the dismissal of performance based education as a viable route to the training of better teachers before it has received a fair test. The Task Force believes that, although performance based teacher education must ultimately stand or fall on the basis of summative evaluation which compares teachers trained by performance based programs with those who graduated from other kinds of programs, such summative evaluation should not and cannot be undertaken until performance based programs have had the opportunity to mature.

This does not, however, mean that attempts at program evaluation should be postponed until PBTE programs have achieved some optimum stage of development at some time in the indeterminate future. On the contrary, there is a real and present need for systematic efforts at formative evaluation—evaluation which collects data on the success and efficiency of many aspects of program operation, data which are then fed back into the program itself in such a way as to allow program decision makers to optimally shape the course of the program. It is the position of the Task Force that performance based teacher education programs should be judged today, not so much by some absolute standard of accomplishment (summative evaluation), as by the relationship between their formative evaluation systems and their decision making structures. Given the state of the art, the ideal evaluation system is one which constantly provides program management with objective, reliable data on all aspects of program operation; conversely, the ideal program manager is one who bases his decisions on data obtained from just such a formative evaluation system.

Such a formative evaluation system should collect data on several aspects of program operation. Of central importance is the evaluation of program attainments (e.g., How satisfactory are the lists of competencies developed? Have competencies and training procedures been validated against a student achievement criterion? Can teachers graduating from the program actually perform the competencies which were specified as objectives? Does the classroom behavior of program graduates differ meaningfully from that exhibited by graduates from other programs?). But program evaluation for today's world must also include such process variables as the adequacy of the management system which supports the program, the locus of program decision making, the flow of communications within the program, and other factors which together ultimately determine the success or failure of the program.

Despite the importance of process variables in program evaluation, however, the crucial question for research remains the selection of adequate criteria of program success. This question is not as simple as it might appear at first glance, for the success of a teacher education program can be judged at many different levels. We might, for example, consider a program

BEST COPY AVAILABLE

successful only if the students of teachers graduated from the program achieved specific levels of academic success. At the other extreme, a program might claim success on the basis of data indicating that its graduates understand the principles of effective teaching and have a good basic grasp of their own subject areas. An effective formative evaluation strategy, however, samples program success on as many levels as possible. This approach permits the evaluator to be highly specific in the delineation of the program's strengths and weaknesses.

The Task Force believes that the criterion levels defined by Turner[1] provide a valuable heuristic device for this discussion. We basically contend that the appropriate level for the evaluation of a competency based teacher education program should be Criterion Level Three; a program should be considered successful to the extent to which its graduates exhibit the competencies for which they were trained.

While Level Three should be the central focus of an evaluative strategy, other levels should also be considered. Individual instructional modules, for example, should be evaluated at Criterion Levels Four through Six; obviously, trainees should be able to demonstrate understanding of specific competencies, demonstrate the performance of these competencies, and use them in front of live students. In a sense, these lower criterion levels might be considered necessary but not sufficient conditions for effective teaching; a teacher who does not understand and cannot demonstrate a competency in a limited situation can hardly be expected to use it effectively in a real classroom situation.

Finally, while it would be unreasonable to hold teacher education programs accountable for the academic achievement of the students of teachers who graduated from the program, a comprehensive formative evaluation strategy should also include research designed to assess the relationship between the competencies covered by the program and student achievement. Program decision makers and evaluators must know how well the teachers whom they trained are doing with their own students, for without this information they cannot adequately evaluate the competencies which are the basis of their programs. They also need to know the extent to which teachers trained by their program demonstrate under actual classroom conditions the competencies which they acquired in the program.

What the Task Force is suggesting, then, is that program evaluation for today's performance based teacher education programs should possess three characteristics: (a) it should be formative rather than summative; that is, it should be used not to judge the ultimate success of the program against some absolute criterion, but rather to help program managers shape the program toward ultimate success; (b) program evaluation should have a systems orientation rather than a goal orientation--evaluation should not end with the determination of whether or not program objectives have been achieved, but should also address itself to the processes by which the system functions; and (c) program evaluation should be supplemented by program validation—in addition to determining the extent to which the system reaches its goals, a research effort should be aimed at determining the validity of those goals in terms of the relationship between the competencies developed by the program and student growth.

BEST COPY AVAILABLE

FOOTNOTE

[illegible] L. Turner, "Levels of Criteria," in [illegible] ed., *The Power of Competency Based Teacher Education* (Boston, Mass.: Allyn and Bacon, 1972).

BEST COPY AVAILABLE

13 Criteria for Describing and Assessing Competency Based Programs

W. Robert Houston, J. Bruce Burke,
Charles E. Johnson, John R. Hansen

During the past few years competency based education programs and projects have proliferated extensively. Some closely reflect the criteria set forth by Elam.[1] Other programs claiming to be CBE appear to be only slight modifications of more conventional approaches. Surveys[2] of CBE practices reflect considerable activity, but the quality appears to vary greatly. Some have simply translated their old programs into the "form" of CBE, while others have diligently applied CBE principles.

But both claim to be CBE. Attempting to describe or compare results of such programs is an almost impossible task.

Beginning in 1973, the Consortium of CBE Centers[3] began a project to describe the various dimensions of CBE as reflected in operating programs. Such a tool could provide the basis for activities such as:

(1) Surveys of CBE activity

(2) Self-assessment of intent and progress by CBE programs

(3) Planning a document to be used by professional preparation programs

(4) Discussion device for considering the function and value of various criteria

(5) Research in institutional change, programmatic strategies, and organizational constructs

The following criteria serve as the basis for this effort. While still regarded

BEST COPY AVAILABLE

as a working list. It represents the third major revision and considerable debate over the past year by a wide range of persons. In its final form, to be published by the Multi-State Consortium in the fall, 1978, each criterion will be supported by a set of indicators and program descriptors.

The purpose of this is to provide another tool in the improvement of professional education programs. Feedback from readers relative to these criteria and to the finished document will be appreciated.

Criteria for Assessing the Degree to Which Professional Preparation Programs Are Competency Based

Competency Specifications

1. Competency statements are specified and revised based upon an analysis of job definition and a theoretical formulation of professional responsibilities.
2. Competency statements describe outcomes expected from the performance of profession-related functions, or those knowledges, skills, and attitudes thought to be essential to the performance of those functions.
3. Competency statements facilitate criterion-referenced assessment.
4. Competencies are treated as tentative assessments of professional effectiveness, and subjected to continual validation procedures.
5. Competencies are specified and made public prior to instruction.
6. Learners completing the CBE program demonstrate a wide range of competency profiles.

Instruction

7. The instructional program is derived from and linked to specified competencies.
8. Instruction which supports competency development is organized into units of manageable size.
9. Instruction is organized and constituted so as to accommodate learner style, sequence preference, pacing, and perceived needs.
10. Learner progress is determined by demonstrated competency.
11. The extent of learner's progress in demonstrating competencies is made known to him throughout the program.
12. Instructional specifications are reviewed and revised based on learner feedback data.

Assessment

13. Competency measures are validly related to competency statements.
14. Competency measures are specific, realistic, and sensitive to nuance.
    14.1 Procedures for measuring competency demonstration assure

BEST COPY AVAILABLE

[illegible]

14.3 Competency measures allow for the influence of setting variables upon performance.

15. Competency measures discriminate on the basis of standards set for competency demonstration.

16. Data provided by competency measures are manageable and useful in decision-making.

17. Assessment procedures and criteria are described and made public prior to instruction.

Governance and Management

18. Statements of policy exist that dictate in broad outline the overall structure, content, operation and resource base of the program, including the teaching competencies to be demonstrated by and from the program.

19. Formally recognized procedures and mechanisms exist for arriving at policy decisions.

19.1 A formally recognized policy making (governing) body exists for the program.

19.2 All institutions, agencies, organizations, and groups participating in the program are represented in policy decisions that affect the program.

19.3 Policy decisions are supported by, and made after consideration of, data on program effectiveness and resources required.

20. Management functions, responsibilities, procedures, and mechanisms are clearly defined and made explicit.

20.1 Management decisions reflect stated program philosophy and policy.

20.2 The identified professional with responsibility for decision has authority and resources to implement the decision.

20.3 Program operations are designed to model the characteristics desired of schools and classrooms in which program graduates will teach.

20.3a Job definitions, staff selections, and responsibility assignments are known to the management functions that exist.

20.4 Formally recognized procedures and mechanisms exist for arriving at the various levels of program management decisions.

Staff Development

21. Program staff attempt to model the attitudes and behaviors desired of students in the program.

22. Provisions are made for staff orientation, development, and improvement.

PART IV

EFFECTIVENESS OF INSTRUCTIONAL STRATEGIES

14 Evaluating Ways to Help Teachers to Behave Desirably

BEST COPY AVAILABLE

Research on Teacher Education

But now let me turn to the second kind of relationship—that between teacher education procedures and methods, on the one hand, and teacher competencies, on the other. How can we know when a teacher education method is producing a teacher competency, or a desirable kind of teacher behavior, or the use of good models or methods of teaching, or what Gagne (1973) recently identified as more important, in his opinion, than any of these, namely, "events that support learning processes"? In short, how can we evaluate the instructional strategies that are used in teacher education programs?

"Knowing That" and "Knowing How"

My answer is based in part on a distinction between two kinds of competencies. This is the distinction first made, as far as I know, by Ryle (1949)[10] and since then further analyzed by Roland (1961),[11] among others—the distinction between "knowing that" and "knowing how." The first kind of knowledge, "knowing that," refers to the ability to state factual propositions, such as knowing that Skinner coined the term "operant" or that teacher education is a controversial subject. "Knowing how," on the other hand, refers to skills or operations, and Roland added the further qualification that the distinguishing feature of such skills is that they imply learning through practice. As she put it, "... Jones could not know how to swim or speak French unless he had at some time practiced swimming or tried to speak French."[12]

My purpose in referring to these distinctions is to delimit the kinds of outcomes or objectives of teacher education with which I shall deal. In some conceptions of competency, or performance, based teacher education, all outcomes are embraced. Such conceptions include the teacher's general or liberal education, the teacher's knowledge of the subject he will teach, the teacher's knowledge about curriculum and instruction, the teacher's knowledge about the social, historical, and philosophical foundations of education, and finally, the teacher's knowledge of how to teach.

I place most of these kinds of knowledge in Ryle's "knowledge that" category. I also prefer to omit from my discussion any consideration of ways of helping teachers to acquire such knowledge and ways of evaluating techniques of giving teachers such knowledge. The reason for restricting discussion in this way is twofold: First, ways of inculcating and evaluating such knowledge are relatively well established. Teacher education's main weaknesses, in my opinion, have not been located in this realm. When it comes to giving teachers knowledge of this kind, we have done about as well or as poorly as the rest of higher education. Second, many of these kinds of knowledge are provided by parts of the university outside of schools or colleges of education, and professional educators can impose performance based approaches outside their own domains only at the risk of arousing considerable academic opposition.[13] My inclination is to try to establish the value of performance-based teacher education in those domains for

BEST COPY AVAILABLE

which professional educators are primarily responsible. And, of these, the one that provides the greatest room for improvement and the greatest opportunity is the realm of "knowledge how"—the realm of skills and habits in teaching. This is the realm in which we include what the teacher does in making decisions about how to teach and how he then uses the skills necessary to carry out those decisions. "Knowing that" is illustrated by the student's ability to state and even explain Archimedes' and Newton's laws. "Knowing how" is illustrated by the student's swimming. Although swimming is explained by Archimedes' and Newton's laws, knowing those laws is one thing, while knowing how to swim is another. The two kinds of knowledge are independent of one another. Similarly, knowing the laws of learning and being able to expound operant conditioning or cognitive theory is one thing, but even though those theories deal with what makes certain kinds of teaching possible, knowledge of those theories is not sufficient to enable one to teach.

It is this realization that has led teacher educators to include student teaching, or some other kind of practice, in their programs. Teacher educators may not have been aware of Ryle's distinction between knowing that and knowing how, or the many subtle elaborations of that distinction that have been formulated, but they have acted upon that distinction. It may be that distinction which most clearly lays the basis for our whole concern that teacher education be performance based. We want teachers to know *how* to teach, not merely to know *that* certain things about teaching are true. This realization has also led to all the attempts of the last decade to improve the effects of practice for prospective teachers—attempts in the form of microteaching, minicourses, and the hundreds of teacher training products that have been developed in recent years.

How to Evaluate Teacher Training

How should we evaluate the strategies that we develop to improve the degree to which teachers behave desirably? In dealing with this topic, we must set up hypotheses about desirable ways of teaching. Furthermore, we must set up hypotheses about strategies for getting teachers to behave in these desirable ways. Both of these kinds of hypotheses have frequently been made in the developmental work of the last few years.

Teacher Training Products

That inference seems justified by the plethora of teacher training products that have been developed, identified, and catalogued by organizations in Florida, at the University of Houston, and at Stanford University. In the Stanford catalog, which had the benefit of being based upon several predecessors developed elsewhere, more than 600 such products have been described in terms of more than 100 descriptors, each of which has provided the basis for an index of the catalog. The Stanford catalog is stored on computer tape, and information about teacher training products with any

BEST COPY AVAILABLE

given characteristic or set of characteristics can be readily obtained. Thus products intended to improve (a) question asking skills in (b) secondary school (c) mathematics teaching, can be identified quickly.

One descriptor in the Stanford Catalog of Teacher Training Products was "References describing field test results." Of the 657 products described in the catalog, only 87, or 13 percent, came with references to published field test results. For another 83, or 13 percent, of the products, the source catalog stated only that the product had been "field tested by the developer." For 22, or 3 percent of the products, reference was made to unpublished data or an unpublished report. When the developers were asked for information by mail, positive results were claimed for 7, or about 1 percent, of the products. For 12 of the products, the developer indicated that field testing was planned or in progress. For 77 products, no field testing had been performed, or it was indicated that the results were available only to the developer of the product. Finally, for 369, or 56 percent, of the products, no information concerning field testing was available in any form.

Thus far we have not determined what kind of information was provided in the 87 references to published field tests. Nor have we ascertained the bases for the developers' claims of positive results for seven of the products. Nonetheless, even without further investigation and analysis, the broad outlines of the field test information indicate that data are available in too few cases, in too little detail, and in a form too unavailable for evaluation by potential users or critics of the product.

The field testing of teacher training products can be expensive and difficult. In many cases, it may not seem important enough to justify the cost in view of the nature of the product and past experience with products of its kind, such as manuals or programmed instructional materials. The field testing of a product for validation can cost much more than the development of the product itself.

Yet it must be clear that the present state of affairs should not be allowed to continue much longer. Standards should be developed in this field comparable to the ethical and professional standards that have been promulgated by the American Psychological Association (1966) for psychological tests and manuals.[14]

Evaluative Experimentation

What form should such standards take? How should strategies for inculcating teaching competencies be evaluated? To a considerable degree, the broad outlines of answers to these questions are already available. In general, we can evaluate such products or strategies by using them as independent variables, or treatments, in experiments or quasi-experiments. Teachers or prospective teachers given training with the procedure or product should be compared with teachers or prospective teachers not given such training. Designs for such experiments were clearly set forth by Campbell and Stanley (1963)[15] and have recently been treated with specific reference to the problem of evaluating teacher training products by Okey and Ciesla (1973).[16] The latter exposition focused particularly on three

BEST COPY AVAILABLE

quasi-experimental designs because of the presumed difficulty, in research with teachers, of obtaining control groups of teachers. These three quasi-experimental designs are the time-series design, the equivalent time-samples design, and the posttest-only control group design.

These matters of design are well known. And the latter three designs have been well illustrated in their application by Okey and Cicsla to the evaluation of an actual teacher training product, namely, their own *Teaching for Mastery* self-instructional program. So I shall not deal further with these aspects of the problem of evaluating ways to get teachers to behave desirably. Let me turn instead to three aspects of the evaluation problem that are, as I see it, somewhat less well known and understood. These are (a) the use of criteria at various levels on the continuum of immediacy-ultimacy, (b) the use of unobtrusive measures, and (c) the use of factorial designs.

My first point is aimed against a dogma which appeared as a reaction against the nonuse of student achievement as a criterion in studies of teaching. Until the 1950's, criteria of teaching effectiveness usually took the form of ratings by supervisors, principals, experts, or students. The bibliography by Domas and Tiedeman annotated more than 1,000 studies, of which only a few used student achievement as the dependent variable.[17] In 1954, Ackerman published a review of the few available studies that related teacher competence to pupil change,[18] and he was followed in 1956 by Mitzel and Gross with a similar review.[19] Since then, perhaps 100 studies, mostly correlational studies, of relationships of teacher and teaching variables to student achievement have appeared and have been reviewed in the volumes by Rosenshine and by Dunkin and Biddle (1974). More studies of this kind are needed—many more.

But my point is that these are not the only kinds of studies worth doing. Not everything desirable about teaching or teacher education needs to be justified on the basis of its relationship with student achievement. Achievement is at one end of a causal chain, or at one corner of a causal network. Many desirable factors in teaching or teacher education may be too remote from that final link in the chain to have a discernible effect upon it. They are desirable, nonetheless, if they can be shown to have good effects on links closer to the treatment, or independent variable, that we are evaluating. In medicine, not every feature of the physician's technique is validated on the basis of its correlation with patient recovery or longevity. It would be difficult to find such a correlation for every detail in methods of taking blood pressure, using an opthalmoscope, swabbing a throat, applying a plaster cast around a broken leg, sewing a suture, or massaging a muscle. The criteria for the effectiveness of such competencies on the part of physicians are not patient longevity, or even patient recovery, but rather outcomes much closer in the causal chain to the particular competency in medical practice. Does the throat swab cover the tissue it should cover? Does the plaster cast endure long enough without being too heavy? Does the suture hold the wound together and leave a minimal scar? Does the massage restore circulation and relieve an ache?

By the same logic, we should ask of a technique of teacher education not only, and not always, whether it results in improved student achievement, retention, and even transfer of cognitive and social-emotional objectives.

BEST COPY AVAILABLE

We should ask first, perhaps, whether the teachers being educated with the technique have some respect and some liking for it. Unless they take some pleasure in the technique, intellectual or social pleasure, or both, the teacher training method will probably not be effective, because it will not be used. Second, we should ask whether the teacher training method has the intended effects on teacher behavior or performance, either in the interactive situation with students or in the teacher's preparation, planning, and decision making before he meets his students. If teacher attitude and teacher behavior are influenced as intended, we are justified in regarding the teacher training as validated in an important sense. We should not insist that every single procedure or product for training teachers be revalidated by its effects on student achievement.

Further, just as criteria come in hierarchies and not only those at the top are worth using, so settings for training teachers come in a hierarchy of realism. We can go from training animals to tutoring individual students in the laboratory to teaching groups of 5 to 10 students in a microteaching setting, and perhaps other intermediate settings, before we come to real classes in real schools. Not every product or procedure should be required to prove its worth on the basis of its effects in actual classrooms. Here again we should try to make reasonable inferences in a graduated sequence.

In short, proximate criteria of all sorts are justified. We are likely to make many errors in which we reject good methods as ineffective, if we always demand that every teacher education procedure or teacher training product result directly in improved student achievement.

Let me turn now to the use of unobtrusive measures. The term was originated by Webb and others in a 1966 volume by that title. It refers to "social science research data not obtained by interview or questionnaire." Research on teaching has not been as prone as other kinds of behavioral science research to depend entirely on questionnaires and interviews. The observation of social interaction, especially in the classroom, has been more highly developed in research on teaching than in other fields of the behavioral sciences. Nonetheless there is need for additional ways of measuring the outcomes and processes of teaching and of teacher training, especially in the non-cognitive domains. We need ways that are not subject to the biases and weaknesses of tests, questionnaires, and interviews. Webb and his co-authors gave a bountiful of examples of unobtrusive measures, which they classified as those depending on (a) physical traces of erosion and accretion, (b) the use of archives, including running records of various kinds, as well as episodic and private records, and (c) observation, either simple or contrived.

Of these, research on the evaluation of teacher training strategies has already used much simple observation and some contrived observation, depending on hidden hardware and controls, such as remote-controlled videotape recorders. But it has made too little use of the other kinds of unobtrusive measures. I can offer only a few examples of what I have in mind. If a helicopter flies over the playground of a multiracial school and photographs the upturned faces, the picture will yield an unobtrusive measure of racial integration in the playground. (I am indebted to Harold B. Gerard for this example.) Similarly, if the use of a minicourse is made

BEST COPY AVAILABLE

completely voluntary, how many teacher trainees even look into it, and how many voluntarily complete it? If a teacher education program is changed, what happens to the characteristics of those who apply for admission to it—such characteristics as prior grade-point average, sex, race, ethnic and socioeconomic background, and scholastic aptitude? What happens to the longevity or the profession of its graduates? What happens to the incidence of broken windows and other kinds of vandalism if teachers are trained in methods of enhancing the students' sense that they control their own fate? What can we tell about a reading teacher's effectiveness from an unobtrusive count of the number of books checked out of the library by his students on their own initiative? What can we tell about the interest of a teacher's choice of topics, before and after the teacher has been given a relevant microcourse, by unobtrusively listening to students' conversations as they leave the class or ride home on the school bus? What can we tell about the effectiveness of a teacher's unit on nutrition by watching what his students buy in the school's cafeteria, store, and vending machines? My examples may be inadequate, but the point remains. We could benefit from greater ingenuity in evaluating teacher training strategies through supplementing present methods with ethically but effectively used unobtrusive measures.

Finally, I turn to the question of how to evaluate a teacher training system as a whole while also evaluating its parts. This kind of problem arises when we want to know not only whether the system or program as a whole has made a difference, but whether each of its components carries its weight and even whether the effectiveness of the whole system is greater, through synergistic interaction among the parts, than the sum of the effects of the parts. The classic solution to this problem is factorial design. For example, if there are two microcourses, A and B, there can be four treatment groups, namely teachers who take A only, those who take B only, those who take both A and B, and those who take neither. But if there are three microcourses, or other components, in a training system, the number of possible treatment groups becomes 8, with 4 components, 16, and so on. From such a design one can learn about the so-called main effects, that is, the effects of each component by itself, and also about interaction effects, whereby two or more components together have an effect different from the mere sum of their separate effects.

For complex teacher training systems, complete factorial designs are unworkable simply because they require too many teachers. Further, they would yield much unpromising information, such as information about interaction effects that are highly unlikely to occur or too hard to interpret if they do occur. For these reasons, statisticians have invented confounded factorial designs, fractional factorial designs, and incomplete block designs. The details of these designs are beyond the scope of this paper. But research on teacher education should eventually make use of these designs, which sacrifice some of the combinations or systematic ways and some of the results on interactions, while becoming much more manageable.

I know of only one example of this kind of research. It dealt with an elemental analysis of the Keller method of instruction. Calhoun determined whether six elements of the Keller plan made a difference in student

[illegible]

[illegible]

[illegible]

15 Task Force on Evaluating Instructional Strategies

16 Priorities and the Nature of Evidence

BEST COPY AVAILABLE

and business and industry regarding similar systems, e.g., storage and retrieval of airlines information; human relations vis à vis the community in education and human relations vis à vis the community in urban planning

11. A plan for more effective sharing of successes between education and business and industry regarding the relationship between their respective "Storehouses" and policy decision making

Secondary

1. More effective systems of storage and retrieval of validated products and processes
2. More effective systems of validating products and processes
3. Better understanding of research and development needs
4. Less duplication and overlap of research and development efforts
5. More systematic relationship between theory and practice
6. Development of levels of evidence regarding educational success

The Nature of the National Storehouse

What It Is

The National Storehouse is a collection of evidence—materials, programs, ideas, concepts, and practices—of what works in education today. The Storehouse exists. What does not exist is an organizational framework that could point up gaps in what we know, stimulate efforts to close them up, and provide a typology that would embrace new, as yet unthought of problems. The missing organizational framework would make the evidence easily accessible to users and would provide information on validation.

What It Is For

The National Storehouse would be a user or consumer based index with entries classified by topics focused on use or intended users. It would therefore be indexed in terms of problems or tasks an educator must cope with in order to be effective:

- the practitioner identifies a teacher competency for which he wants to provide training, so turns to the National Storehouse for a search of what is available
- the policy maker is charged with formulating a new piece of legislation and needs information about it (e.g., reading program), so turns to the National Storehouse for research evidence and a bank of knowledgeable people
- the high school teacher of English wants training in how to cope with student response to literature, so enters the system seeking English, high school, student response—and the level of evidence desired. The computer search would identify the appropriate alternative materials relating to the problems as described

BEST COPY AVAILABLE

... the policy maker ... a school board, interested in improving its elementary staff development program ... could enter the Storehouse by identifying school personnel, teacher training, elementary school, in-service, in order to see if materials exist that address this question.

In organizing the National Storehouse, therefore, we address the question, how will it assist

teachers
teacher educators
legislators, national policy makers
researchers
educational consumers
educational administrators
lay administrators

What It Does

The National Storehouse identifies what we have and, as an immediate outgrowth, identifies what we need. The National Storehouse, therefore, shaped by need (or the consumer), identifies for the potential user the origins (theory, experience, hunch, invention) of the evidence, then it explains its current state of development (idea, pilot test, programmatic test, experimental test), its target, its relevance to broad categories of problems (policy decisions, research and development programs, practical application in teacher training, and organization and administration of schooling).

Organized so it can answer the question, "What do we know about X?" and open-ended as well, the National Storehouse could classify evidence in two ways:

1) by the categories of information a user desires;
2) different conceptions or categories of educational problems. This would permit the user to employ his/her own search strategy and also invite him to look at the items in a variety of different ways.

Such an index could be constructed logically, based on some kind of analysis of problem areas. In a word, the system should be developed deductively from what a consensus of practitioners and laymen say we need, not inductively from what we know we have—although the latter would also be included.

Retrieval and Dissemination

A user- or needs-oriented National Storehouse must assure easy access to practitioners and policy makers.

The first step toward easy retrieval might be the organization of the index of storehouse materials into a simple directory. Contents might be divided into: (1) printed materials, (2) audio-visual materials, (3) ideas, (4) people and places. Listings should be short—product, contact name and address, one-line summary, extent of validation, cost. The directory should be set up as a loose-leaf binder so that pages may be printed and distributed for insert. It should also be free, or the cost should be nominal, covering only

making dependent. Obviously, then, the directory must be funded by Federal or foundation moneys. This foundation or Federal grant should be large enough to permit some subsidization of the producers of the items listed in the catalog while producers seek ways to bring down production costs.

Unless materials costs are brought down, the many access suggestions offered for practitioners will not be accomplished.

But to get that directory into the hands of the practitioners will require a major dissemination program. Such a program should be launched by a communications expert who is competent to utilize all media approaches to the problem. Such approaches would include such standard methods as direct mail notifications to local and state education agencies, to associations and unions, to teacher centers, and to the various national, state, and local professional education associations. But the dissemination program should go much further. Exposure should mainly involve public service time and space in the education press (national, state, and local), the education trade and professional press, local, closed-circuit and cable radio and television, local and state conferences, displays at National conferences, and so on.

The message should be simple. American education is probably the most experienced and diverse in the world. And much of what has been learned about the way to do things better is available for educational program planners.

Costs for such a program should be explored. A follow-up to (but also simultaneous with) the dissemination program might be the use of trained leaders who could go out on request to work with local teachers on use of the materials and methods. The communications expert could make suggestions on how best to coordinate this. But unless state and local education agencies—or a foundation—can underwrite such a program, costs might well be prohibitive.

Quality Control

But a mechanism might be developed to inform users about the quality of the evidence.

Evidence should be thought of as a device for either hypothesis testing or decision making. Evaluation criteria differ, depending on which of these two purposes the evidence is to serve. Evidence is generally "harder" for hypothesis testing than for decision making, although this may account for bad decision making.

The second factor determining the quality of evidence is the stage of development of the idea, program, or product. Different evidence is expected for a product in its earlier stages than in its "finished" stages. Thirdly, things that work—hard evidence—within the context of an integrated and sequenced program may not work when taken out of that context.

Rather than search for "hard" evidence, the functions of rating might better be served by classifying items according to where they are in a sequence of steps in the research and development process, rather than rating them. The sequence might run from armchair ideas, entirely untested, through ones tried out in laboratory experiments, in simulation or micro-

teaching settings, to those best tested and ready for dissemination. It should be possible to define in an quite practicable and precise a piece of evidence set in this sequence should be of much to useful to know as how it was rated by somebody. A this kind of rating system was employed. It would be possible to enter the Storehouse asking for the level of develop- ment (e.g. how much the Storehouse had in that subject at a particular purpose that had been thoroughly field tested). A display could be called for which would show such a quality control process by area, and it would be easy to see where weaknesses and strengths lay - a useful tool for planning for communicating with legislators, etc.

How To Organize It

A number of models have been proposed, and they are by no means mutually exclusive.

(1) Use a

Develop information sources:

1. Gather available materials, including ideas under four broad headings: (a) teaching skill or performance, (b) curriculum materials and design by intended outcomes, (c) training of educational personnel, (d) organization of instruction and schooling.

2. Prepare for each item a developmental history - who developed the product or idea, its theoretical or other sources, its use, its current status.

3. Classify each item by the criterion of its effectiveness that the user has selected and by other criteria to which it might be relevant.

4. Gather evidence from the developer and other users about evidence relative to effectiveness with respect to the criteria.

5. Sort items by their relevance to policy, research and development, or practice.

(2) Develop the tagging system so that anyone who wants to search the system can do so:

1. invite groups of potential users to ask questions of the system.

2. Tag items by question(s).

3. Identify unanswerable questions.

4. Determine whether the question is unanswerable because of the classification system or because there is no information in the system.

(3) Develop validation of system:

1. Invite conceptualizers of problems to ask questions of the system (e.g., Jencks asks the system for information on size as effectiveness; Coleman asks the system for information on the effects of desegregation).

2. Invite a group of people to use it then ask them to answer (e.g. ask 4 questions in a period answer).

Ultimately, evidence might be collected to examine goal and usage of BCP development.

Goal

Produce Significant Study Items

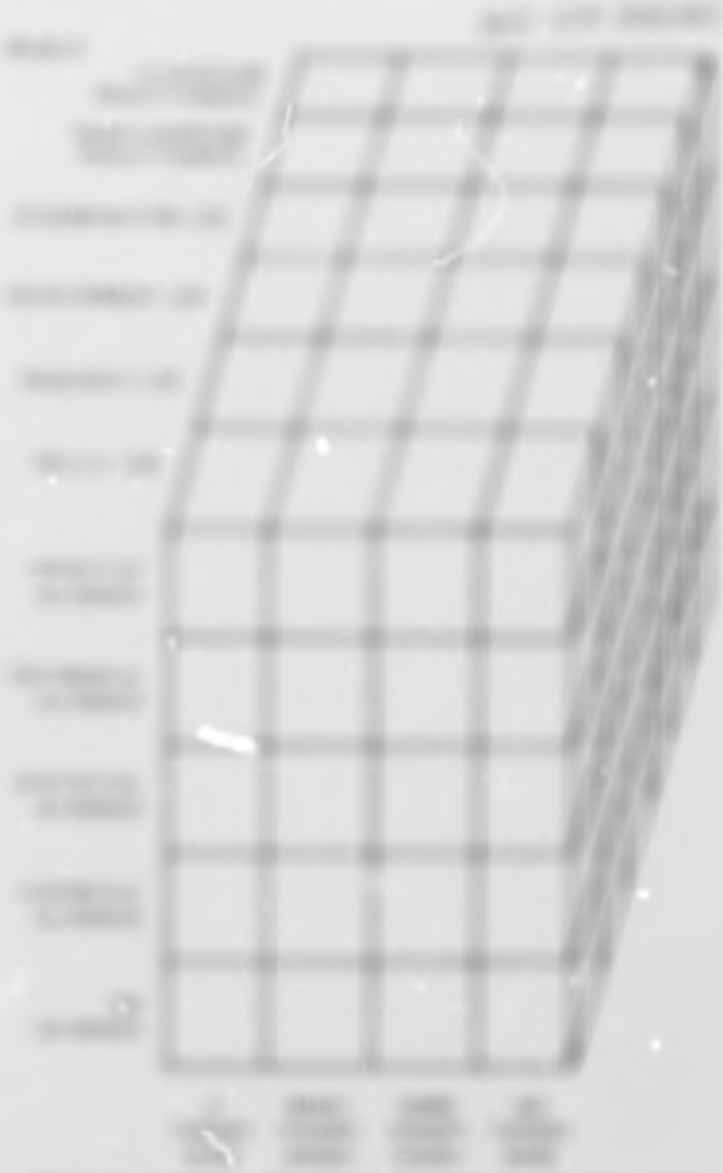
Area	Development			

Model of

1. Develop user-user categories:

Primary User	Primary Use
Teachers	Education of Children at various Levels
Teacher Trainers	Preservice Teacher Education
Teacher Trainers	Inservice Teacher Education
Decision Makers	Educational Decision Making at various Levels
Decision Makers	Policy Decision Making at various Levels
Researchers	Heuristics
Evaluators	Mission Research
General Public	Summaries

Each user or user category would generate to user accounts, though perhaps [illegible] overlapping set of categories. Designed to make sense to those interested in [illegible] for that purpose (e.g. 3/10 left might be of high interest for elementary teachers, moderate interest for teacher trainers, little interest for decision makers, no interest for neurosis researchers, moderate interest for mission researchers, and be of great interest for synthesizers, and of great use in dissemination to the general public).



BEST COPY AVAILABLE

Gage, N. L., et. al. *Computerized Catalogue of Teacher Training Products*. Research and Development Memorandum No. 116. Palo Alto, California: Stanford Center for Research and Development in Teaching, Stanford University, 1974.

Hudgins, Bryce B. *A Catalogue of Concepts in the Pedagogical Domain*. St. Louis, Missouri: Washington University, Graduate Institute of Education, 1974.

Guidelines for Delivery of Development Products (Draft). Washington, D.C.: National Institute of Education, 1974.

Protocol Catalog. Materials for Teacher Education. Tallahassee, Florida: State Department of Education, 1974.

Schmieder, Allen A. and Joyce, Bruce. *The National Storehouse* (for CBE), two versions—full and abridged. Washington, D.C.: USOE, Office of Career Education, Education Systems Development, 1974. (Draft)

Schmieder, Allen A., and Mark, Jori, eds. *The National Storehouse of Educational Materials/Successes*. Washington, D.C.: USOE, Office of Career Education, Education Systems Development, 1974. (Draft)

Sharing Education Success. A Handbook for Validation of Educational Practices. Washington, D. C.: United States Office of Education, National Advisory Council on Supplementary Centers and Services, 1974.

Turner, Richard L. (ed.). *A General Catalog of Teaching Skills*.

The National Storehouse of Educational Materials/Successes: Profile of Some Potential Experience Bases

1. 16,500 school districts
   - 700 "community" schools
   - 325 schools using individualized instruction
   - 4,500 teaching centers
   - x alternative schools
2. 50 states
   - 20 CBE mandates
   - 9 career education mandates
   - 4 teacher center mandates
3. 1,200 teacher training institutions
   - 120 CBE programs
   - 600 teacher centers

The National Storehouse of Educational Materials/Successes: Some Interesting Facts

1. Many powerful educational policy makers do not think that it exists.
2. Many educators feel that it is already too bountiful and needs no further money for development.
3. Many feel that ERIC has the Storehouse well under control.

BEST COPY AVAILABLE

4. Storehouse materials with less validation sell faster.
5. Most new validated materials are either on a shelf somewhere or used in a single classroom or program.
6. Most consumers do not like to go to some "other" place to use catalogs and materials banks/resources—they want them located in their office.
7. The Newsletter is one of the highest demand consumer items in the dissemination business.
8. There is a very thin market for many of the most important materials developed.
9. Regarding product validation—what might be hard evidence to one kind of consumer is useless to another kind!
10. Many validated approaches are born in theses and dissertations—and eventually die there.
11. Most producers are reluctant to share materials (a) which are unvalidated; (b) which might be plagiarized or copyrighted by someone else.
12. Publishers are reluctant to produce high risk materials.
13. All producers are potential consumers and all consumers are potential producers.
14. College and university staff generally feel that they are the only ones sufficiently qualified to produce valid materials for widespread use—school staff generally feel that college and university staff generally don't know enough about the real world to produce valid materials.
15. Materials in the process of "becoming" are usually more related to "up front" research and development people than those that have "already become"—yet dissemination strategies almost always relate to "finished products."

The National Storehouse of Educational Materials/Successes: Contributors

Theodore Andrews
Robert Ardike
Michael Berger
Joel Burdin
Patricia Cabrera
Paul Collins
Ned Flanders
Daniel Ganellis
Carl Grant
John Hansen
Donald Hawkins
Mary Johnson
Bruce Joyce
Susan Klein
Jori Mark
Karl Massanari
Moira Mathieson
Donald Medley
Fred McDonald
Donald Orlosky
Lorraine Poliakoff
Preston Royster
Allen Schmieder
Joseph Schneider
William Smith
Francis Sobol
James Steffensen
William Tikunoff
Wilford Weber
Margaret Wiesender
Philip Winne
Sam Yarger

BEST COPY AVAILABLE

PART V

SOCIAL CONTEXT

BEST COPY AVAILABLE

17 The Politics of Teacher Competence

Harvey B. Scribner
Leonard B. Stevens

Since the matter of politics generally is not raised in most discussions of teacher competence, the title of this paper perhaps requires a brief explanation. "Politics" as used here does not refer to the partisan activities of political parties or the influence that political officeholders may have occasionally or regularly in issues of education. Rather, "politics" refers to the activities—which, though nonpartisan, are keyed to identifiable interests and goals and thus are, in the broad sense of the word, political activities—and pursuits of groups and individuals within the arena of education.

American public education is well known for its unpolitical character. Clearly, neither the public schools nor the public universities developed in a political or social vacuum. It is equally clear that there has been a notable and commendable absence of overt political, partisan control of the nation's systems of public education.

From this, however, one should not assume that educational institutions function without a political atmosphere. They do have and always have had a political atmosphere. Local school boards, for example, typically pride themselves on their independence from city hall. Yet all but a handful of school boards are dependent on political officeholders for raising and appropriating public tax funds for schools. Moreover, every school board to some extent participates in the deployment of fiscal and human resources for reasons that more than occasionally stretch into political or quasipolitical considerations. Similarly, while state departments of education typically work hard to develop an insulation from statehouse influence, rare indeed is the chief school officer who does not know his way around the halls where statutes are written and appropriations are made. Nor can those who man-

BEST COPY AVAILABLE

age the public universities be accused of total innocence of the art of politics. Despite the academy's characteristic disdain for politics, university heads with budgets and construction programs in mind find it necessary to do periodic battle in mending with elected representatives of the public who control the strings to the public purse. Inside the academy more than a small amount of energy is expended on political maneuvering.

In brief, the public education enterprise, despite its unpolitical reputation, is quite political when "politics" is defined in broad terms. Schools and universities affect, and are affected by, external political forces. In addition, education systems, because they are large organizations with important social functions and immense resources, possess a set of internal political dynamics which at times become quite intense. Teachers and professors, to organize for special ends, and they are not disinclined to display their collective power to achieve their ends. Administrators are no different, except that they have the advantage of power which comes from their position. Likewise, school boards and university trustees have political power, and they use it. In addition, they band together in organizations on the regional and national levels, and these organizations themselves wield a certain amount of influence over the nation's education systems. Even researchers, who often are regarded as the most objective and least political of all inhabitants of the education forest, clearly have their own values and distinctive perceptions of what is and what ought to be. And to the extent that researchers channel and shape their research and their conclusions in keeping with their value system, they, too, participate in education politics. "Man is by nature a political animal." Aristotle told us long before politics acquired its undeserved, distasteful connotation. And while he perhaps did not have educators in mind, the observation clearly did not exclude them, as contemporary history of schools and universities has made quite evident.

There was a report in the *New York Times* recently that two public agencies in New York City, the Board of Education and the Board of Higher Education, had met together for the purpose of establishing closer coordination of their respective activities and programs. Now, since both agencies have responsibilities in public education and since both spend public dollars, the effort to coordinate surely is commendable, and the *New York Times* was quick to do precisely this a day or so later on its editorial page. As a kind of footnote to the story, it was reported that once the heads of the two agencies had met, the next thing they did was to meet privately with the heads of unionized teachers and faculty members.

The *New York Times*, it would seem, should have commended the agencies not only for their positive instinct to coordinate, but also for their keen sense of pragmatism.

This is not necessarily to condemn this "touching base" syndrome, though one is left to wonder precisely what private understandings were reached by the agency heads and the staff unions and to wonder as well whether any understandings were reached that might impair the coordination effort in some way. In any event, this incident is a clear example of politics at work. The point is worth making that political activity should be recognized as such by those who participate in it as well as by those who observe it.

[illegible]

It is neither scaremongering on some users to view it not as it an indispensable element in modern public administration, as others sometimes contend.

The issue of teacher competence has other policy questions in education, too, political questions. Anyone who reviews the professional literature on competence and performance in teaching is struck first by the volume of writing on the topic in recent months. The Education Index for one recent 12-month period [illegible] makes reference to more than 200 articles on the education of teachers - no small number when pertaining directly or indirectly to the question of performance or competence. The second striking characteristic of the literature is the influence peddling that currently is going on.

Perhaps 'influence peddling' is too strong a way of putting it. Perhaps one should say that the literature abounds with the work of individuals contributing to a controversial debate. Whatever one wishes to call it, there is no lack of information on where the major constituencies stand in the debate vis à vis the political questions of: a) who is to control the education and licensing of teachers and b) how is competence to be defined. Lines continue to be drawn that make clear who is the behavioralist and who is the humanist; who is the teacher educator and who is the teacher employer; and who is the professional spokesman and who is the public advocate.

One might conclude that education needs a Dr. Kissinger to work out a settlement and achieve a truce - but perhaps not. Perhaps what is needed before the issue can be settled with any degree of finality is a set of guidelines for working toward a settlement. Given the present state of the public schools and the educational concerns of the society to which the schools belong, it would seem that those with a role in the design of policies and programs for teacher competency should keep at least the following general points in mind.

First, one should be aware of special interests, not disguise or ignore them. The policy designer need not necessarily make moral judgments as to 'good guys' and 'bad guys.' But he ought to be aware of the politics of policy. Since policy affects people, people will assert all the influence at their individual and collective command to shape and control policy in its formative stages, so that emerging policy will be in their best interest. Needless to say, the goals of a special interest do not necessarily make for good public policy or for better schools. Often, indeed, the goals of special interests ignore the public interest, though the makers of interest groups might have the public believe otherwise.

In the formulation of performance systems for preparing and licensing teachers, it seems clear that teacher educators, who have long controlled teacher preparation and thus heavily influenced teacher certification, tend to favor quite strongly the retaining of their favored position. This does not mean they are against systems that ostensibly will enhance teaching performance. But it does mean that they probably will have a very special interest in retaining as much control as possible over teacher preparation programs. The teacher educator may typically be biased toward a belief that the center of gravity of teacher education belongs on the university campus, in contrast to the view of many school administrators who favor

The Decision-Making Process in Planning and Implementing Education Programs and What Are the Needs of the Planning Bodies

BEST COPY AVAILABLE

with policy made by more general publics being made to facilitate rather than to constrain the development and implementation of local policy . . . Good policy-making—that leading to reform in the interest of learners—has the following prerequisites:

1. Those to be involved must have organized a policy making body representing all of the constituencies to be affected. The Task Force has named this the Policy Board.

2. What should be done must be determined by all to be involved.

3. Decisions regarding criteria for determining the effectiveness of management and operation must involve all parties.

4. Commitments to shared responsibility in ongoing evaluation must have been established.

5. Agreements on qualification criteria must have been established for selecting and retaining those who will implement policy.[2]

HETFIRE is only one of many educational groups to recognize and attempt to respond to the present educational needs of equal representation and equal participation in education decision making. Within the last 2 years, there have been a significant number of publications in journals and a significant number of conferences which have treated educational reform, governance, and collaboration. By and large, this treatment has been more exhortation than action. Power is not given; it must be taken. It cannot be assumed that because we hold conferences and write articles about governance, parity, and collaboration in regard to educational decision making that these concepts will be implemented. A case in point is a phrase in the Declaration of Independence, nationally adopted almost two hundred years ago. The following words have still not come true for all Americans:

> We hold these truths to be self-evident—that all men are created equal; that they are endowed by their Creator with certain inalienable rights; that among these are life, liberty, and the pursuit of happiness.

In any honest, sincere discussion of the decision-making process in education, attention must also be given to another point—those who participate in the decision making process must represent the racial and cultural diversity of this country. It cannot be taken for granted that just because various role groups are represented, that the racially and culturally diverse populations of this country are also represented—for most often they are not.

BEST COPY AVAILABLE

Most of the types of role groups which have been historically identified as necessary components of any educational decision making body—college administrators, Federal and state officials, leaders in teacher organizations—are those who have been most inaccessible to nonwhite, non-Anglo-Saxon people. In addition, decision making bodies in education have not included all types of role groups; community representatives constitute one such group that has largely been ignored.

Clinical Model of Action

What follows is an outline of a workshop that can be utilized as a method of helping educators to better understand the actual process of planning and implementing CBTE programs that respect multicultural society. The four-part model is based on the premise that actual participation in the process of decision making is as necessary as theory and good intentions.

Part I: The Setting

After being given the rationale for the clinic (see preceding section), the participants will be given a scenario of a college undergoing change. The Scenario is designed to provide the participants with one possible example of a college attempting to embark on a new multicultural CBTE program. The scenario also provides clinic participants with a number of character roles that can be used for actual practice in decision making.

SCENARIO

State College is a multipurpose institution in terms of the diversity of its program of studies and the diversity of its students. It is located in an urban, cosmopolitan city with a heavily diverse population of Anglo-Saxons, Blacks, Orientals, and Puerto-Ricans. Since the Brown *decision in 1954, State College has dedicated itself to meeting the needs of the community by making its educational programs available to all people regardless of race, color, creed, sex, or national origin. It is a 4-year degree-granting institution with master's degrees being offered in Elementary Education, Early Childhood Education, and Social Work. The total educational program is organized into 10 divisions, one of which includes the Division of Teacher Education.*

Teacher Education. *This division trains approximately one-third of the total number of graduates of the college. The students are provided with theory and practice necessary to help them become "proficient" elementary or secondary school teachers.*

The Division of Education includes six departments:

Early Childhood Education

Elementary Education

Secondary Education

BEST COPY AVAILABLE

Health and Physical Education

Audiovisual and Educational Television

Educational Disabilities

Each department is headed by a department chairperson, and the total division is headed by a division chairperson who is a recent addition to the faculty, recommended highly by the dean of the college's academic affairs for his/her competence in affecting change.

Instructional Program. Though the graduates of the Division of Education have been certified and highly recommended in some area schools, there is general dissatisfaction with the teacher education program and its relevance to the social, economic, educational, and cultural aspects of the community. The educational program is noted for its competent instruction in four or five curricular areas. The students are exposed to proficient lecturers and adequate multimedia. The program is designed around separate subject areas, usually characterized by a series of three credit-hour courses with autonomous instructors who focus on their own competence and interest. These curriculum areas are based on independent goals and objectives. There are innovative practices such as microteaching, interaction analysis, a clinical approach to supervision, independent study opportunities, an inner-city tutoring program, and a bilingual tutoring program. The majority of the academic program is campus centered.

Though these innovative practices and traditional elements are yielding positive results, they never fit together in a way which gives maximum payoff to the students or the community. Thus, the division chairperson and the Elementary Education Department chairperson decided to look for ways to put the pieces together into a program design based on careful planning, the immediate and relevant needs of culturally diverse student teachers, and a multicultural school population. They decided that some of their basic assumptions about teacher education, previously based on tradition, needed rethinking in relation to teachers' roles and to what and how we want children to be and to become. After observing several teacher education programs and many models of CBTE, they decided that one way was to begin looking at teacher behaviors, attitudes, and understandings that facilitate pupil growth physically, emotionally, and culturally.

It is evident that State College is involved in a stage of educational reform. The success of this reform effort will depend upon the inclusion of diverse role groups and cultural groups; it will also depend on the interrelationship of those elements essential for planning and implementing an educational program that meets the needs of a multicultural society.

Part II: Planning the Program

After the clinic participants have had an opportunity to read the Scenario, they will be instructed to select volunteers to come before the group and to portray roles either identified in the Scenario or suggested by it, which

BEST COPY AVAILABLE

would be necessary and/or helpful in the initial planning of the program at State College. The clinic coordinator is trying to determine whether the role players will choose to plan their program changes using only the characters associated with the State College or whether they will choose to reach out and involve other community members (i.e., lay individuals and/or groups, school district personnel, labor groups, or possibly state and Federal officials). The coordinator will also be observing whether or not the participants address any of the following questions:

1. Who constituted the program staff established to study the feasibility of developing the program? Were minority group members on this staff?

2. Should an educational needs assessment be made? What educational needs should be considered in planning the program? Was an assessment of minority educational needs made?

3. What ideas or suggestions were solicited concerning multicultural education? Which individuals or groups were asked about this?

4. Should the program seek to change or influence institutional goals and/or objectives? Should goals and objectives relate to the problems of cultural diversity?

The roles that the participants select and the issues they address in the initial planning for the program will be listed. The coordinator will not provide feedback at this time.

Part III: Implementing the Program

The participants are now ready to start implementing a multicultural CBTE program as suggested by the Scenario and as initially planned (see Part II). The role playing will be continued and the participants will be allowed to add or delete roles depending on what they learned from the planning stage or from what they expect they will need for implementing the program.

Again, what is primarily being looked for here by the coordinator is the choice of roles and the issues they address concerning the implementation of the program. Some of the questions the players might raise are the following:

1. What criteria should be used for admission and retention of students in the program? Is there provision for recruitment of minority students?

2. Are teacher competencies specified which relate to understanding cultural diversity? Are attitudinal changes and modifications of

BEST COPY AVAILABLE

teacher behavior included in the CBTE program? Should minority group members be involved in determining what attitudinal changes are needed?

3. What roles should students and faculty have in decision making? Are there provisions for staff development in relating CBTE to needs of multicultural education?

4. Who should select and design instructional materials? Are instructional materials identified according to some criteria relative to multicultural education?

5. What provisions should be made for continuous feedback of information from all participants in the new program to the assessing agent and/or program staff? What plans can be made for program modifications based on feedback? What provisions can be made for feedback and change from minority group members?

The roles adopted by the participants and the issues they address in implementing the program are again noted. No feedback is provided at this time.

Part IV Small Group Discussion

Since only those who have participated in the role playing will have had an active part in the clinic up to this point, all those attending the clinic will now be invited to join the proceedings. This will be accomplished by having at least one of the role players join a group of not more than six other individuals. Each of these minigroups will then be given an adequate amount of time to further refine the issues that were raised by the role players regarding the planning and implementation of the multicultural CBTE program. A brief oral report of these findings will then be made by a spokesperson from each of the groups to the entire gathering.

Part V Feedback and Evaluation

The roles adopted and the issues raised by the participants in Part II and Part III, as well as the additional input provided by the minigroups in Part IV, will be summarized by the clinic coordinator. The coordinator will compare and evaluate the results of the clinic with the findings of a recently completed national survey of CBTE programs.

Report on Survey Results

Very little research has been done to assess the effectiveness of CBTE

BEST COPY AVAILABLE

programs or to determine the characteristics of design, planning, and implementation which result in true effectiveness. The paucity of research has meant that those involved with the development and conducting of the programs are forced to operate in the dark, with few empirically based guidelines to assure that their programs will succeed in attaining the desired ends. To remedy this situation, it is imperative that measures of various program characteristics be developed and examined in the light of their effect upon educational outcomes.

One of the few studies so far conducted toward this end is a survey conducted by Grant and Carlson (1976) which investigated conditions in 61 colleges across America which are in the process of setting up or constructing CBTE programs. One major purpose of this survey was to examine characteristics of these colleges in relation to their effect upon the planning and implementation of the programs and their approaches to multicultural education.

Toward this goal, the Grant and Carlson study employed rigorously constructed scales to measure the amount of input into the programs from students, minority groups, and the community at large; the pluralistic content of the curriculum; responsiveness to minority needs; the amount of attempt devoted to careful planning; and the degree to which administrators perceived their programs as instruments of educational change. One of the more striking findings of this research emerged from a factor analysis of the intercorrelations among these instruments - all of which loaded rather heavily on a single factor. This result indicates a strong degree of cohesiveness among the program characteristics, wherein a program which is strong in any one of the assessed areas will probably be strong in all of the other areas as well. The factor underlying this cohesiveness was tentatively interpreted as quality of program.

When the components of this factor were examined in the context of various characteristics of the institutions supporting the program, such as size, selectivity of admissions policies, library size, the ratio of Ph.D.'s on the faculty, etc., it was discovered that there is virtually no relationship between the success of the program and the quality of the supporting institution. This finding seems to imply that careful attention to the interrelated variables which determine the quality of a CBTE program can result in the development of a good program at any college, almost regardless of the current finances and facilities of the college.

Within the intercorrelations among the scales, it is noteworthy that one particularly strong relationship was found. The extent to which minority group members have input into the program is correlated at r = .76 with the adequacy of the program in meeting minority needs. This relationship is of such magnitude as to suggest that institutions whose programs do not receive input from minority group members will not provide a program adequate to the needs of these people.

In addition to these findings, tabulations of responses to several other items in the questionnaire provided results of interest to the present discussion.

1. A large number of the colleges and universities surveyed did not

[redacted] and criteria for determining who should [redacted] in the initial planning.

3. Over a third of the institutions reported no community involvement in planning those aspects of the CETE program. Those community representatives who were [redacted] were overwhelmingly [redacted].

4. In a large percentage of the programs, minority group members were not involved in determining proposed curricular changes and modifications of teacher behaviors.

5. Only one-half of the institutions have procedures for the recruitment of minority group students into the CETE program.

6. The majority of the available instructional materials were not identified according to criteria relevant to multicultural education.

It is clear, however, that the Grant and Cartoon study represents only an exploratory venture into the area of open questions about curricular CETE. Extensive further investigation is an absolute necessity if educators are ever to produce teacher educators with any confidence that they are aimed at the goals they set out to reach.

Summary

Planning and implementing an educational program that reflects and responds to the multicultural population of our society should be an exciting, challenging endeavor. Indeed, if this idea could be realized, we would have taken an affirmative step toward rectifying a problem which now denies many of our citizens their fundamental right to representation in any decision-making process that seriously affects the quality of their lives. Racism is still among us, the keeper of the gate, consistently. Most programs overlay concepts such as plans and pluralism under the institutional rug. It will take courageous and dedicated individuals to implement such a program because of the overt and covert racism that permeates every aspect of our society. Nevertheless, this does not mean that such a program should not be attempted and cannot be developed.

Finally, it is interesting to note, and is worth repeating again, that one significant finding of this study is that if educational programs are to respond to the needs of minority group people, the programs must have input from minority group people. If we are serious about having a program that responds to our diverse society, we must not [redacted] and know we can get by with a token individual or two we cannot. To reform an educational program to meet the challenges and needs of our diverse society, we must have individuals from different racial and role groups fully participating in the decision-making process. To reject this idea is to further perpetuate racism.

PART VI

INSTITUTIONAL CHANGE

19 Some Unorthodox Views on Institutional Change

BEST COPY AVAILABLE

20 Task Force on Institutional Change

James Dyer
Dell Felder

This task force consisted of from 12 to 15 members (at different times) and decided at the outset to discuss and develop the research issue as a single group. The chairman, Dell Felder, introduced the game play and suggested that the context be broadened to include common issues that influence change in all sizes and types of institutions.

It appeared that two assumptions were made generally by the group (though not actually verbalized until much later). One was that the institution that was referred to was, in fact, the institution of higher education, and the second, that the change referred to was the implementation of CBTE.

Introduction around the group disclosed a variety of representatives accounting for all critical change interest groups with the exception of the professional teacher organizations. The brainstorming technique was used to bring forth a flood of issues, primary areas of interest being increased costs, professional autonomy, administration involvement, and the general relationship currently existing between the university, public schools, professional teacher organizations, the community, and the state departments. (Note: these will be herein referred to as the groups.)

There were strong feelings about the magnitude of change to be instituted in the initial stages, the commitment of the various authority figures, and the fact that change, as such, could never be "rammed down anyone's throat." Many of the group continued to clamor for "more proof" before change, though others recognized the need for establishing valid research in actual experimental programs before such proof could be forthcoming.

The recommendations for research issues and their group-assigned

BEST COPY AVAILABLE

priorities are as follows:

1. What, specifically, can IHE & other groups learn from each other?
2. Should research efforts be shared by the five groups?
3. What are the prerequisites for collaboration between all five groups?
4. Is a new professional school of education desirable?
5. What are the emerging functions of professional schools of education?
6. Are pioneer (pilot) programs desirable?
7. How do you develop a pioneer (pilot) program to influence the entire system?
8. How is a successful prototype program implemented as *the* program?
9. How do (job) supply and demand influence institutional change?
10. How can threat to faculty and professional autonomy be avoided?
11. What rationale is there for nonchange in all five groups?
12. Is the traditional reward system really a major influence on change?
13. What does institutional change include?
14. Do institutional histories and culture make a difference in developing change?
15. Do educators believe teacher education needs to be changed?
16. Can the Teacher Education Center help develop institutional change?
17. Should change (CBTE) be limited to undergraduate areas?

Feeling was strong that the research personnel should include the interested parties, and that all prior related studies be consolidated and promulgated as soon as possible.

BEST COPY AVAILABLE

21 Model for Developing a State System for Managing Performance Programs*

Robert C. Burkhart

A

Overview

The results of a study completed by Dr. Robert C. Burkhart, President of Educational Assessment Systems Corporation, and supported by the New York State Education Department indicate that student teaching evaluation forms in use in many New York State institutions of higher education with approved programs in education consist largely of judgments made on the basis of either subjective requirements or subjective criteria. The results also imply that the evaluation of student teachers emphasizes portions of professional education for which formal instruction is seldom provided.

By supporting this study, the Division of Teacher Education and Certification of the New York State Education Department was attempting to gauge the present state of the evaluation of student teachers in approved programs in education. It was felt that publication of the results of the study would allow teacher preparation institutions to measure their own progress toward the implementation of a competency based teacher education program.

Conclusions drawn from the results of the Burkhart study indicated a

Michael King and Mary Van Ryn, both of the New York State Education Department, assisted in preparing the above article for publication in this volume.

BEST COPY AVAILABLE

significant control between a majority of the student teaching evaluation forms now in use and the assumptions and principles outlined in "The Educational Rights of Students," a statement prepared by Burkhart and Max Van Ryn, Chief of the Bureau of Inservice Education, Division of Teacher Education and Certification, NYSED (see Table 21.1).

To conduct this study, data were gathered from the student teacher evaluation forms used in 89 New York State institutions of higher education with approved programs in education. The data were analyzed to determine the emphasis placed upon various portions of student teaching and the assessment standards employed in making judgments about the performance of the student teachers. A result of the analysis, a "Representative College Student Teacher Evaluation Form" (see Table 21.2), a prototype of the forms now in use, was constructed and analyzed.

The Study

The 89 collected forms were analyzed to determine the total number of discrete evaluations required by all forms. A total of 3,191 descriptive items were identified and 3,015 were coded according to the Educational Assessment Systems Corporation Analysis System. 176 items were discarded because they were too general to be classified.

The EASC Analysis System is a description in list form of the functions required of a student teacher (see Table 21.3). For the purposes of this study the instrument was presumed to be exhaustive. The Analysis System consists of three major areas: The Teacher-Learner Pre-Instructional Functions Area, the Teacher-Learner Instructional Functions Area, and the Teacher-Learner Professional Profile Area. The areas are divided into categories as follows: Pre-Instructional Functions Area, four categories; Instructional Functions Area, four categories; and the Professional Profile Area, two categories. The ten categories are further divided into a total of 67 subcategories. Each subcategory has its own distinctive number. The distinct number of the subcategory was the number used to code the discrete evaluation items discovered in the 89 evaluation forms.

The decisions regarding classification were made by individuals who had received an orientation in the use of the Analysis System. Each item was rated separately by more than one judge and the inter-rater reliability factor was found to exceed .90.

To construct the "Representative College Student Teacher Evaluation Form" (REF), it was necessary to complete two independent analyses of 3,015 items. As a group, the items, coded according to the subcategories in the Analysis System, were tabulated to determine their frequency of use in the 89 forms. While number 23 of the subcategories was not mentioned, the remaining 56 subcategories showed a frequency ranging from 1 to 500 times on the 89 forms.

In the second analysis, each of the 3,015 items was examined to determine the assessment standard employed. The following is a list of the standards found to have been used, the code identification, an explanation and example of each, and the percentage of the total number of items repre-

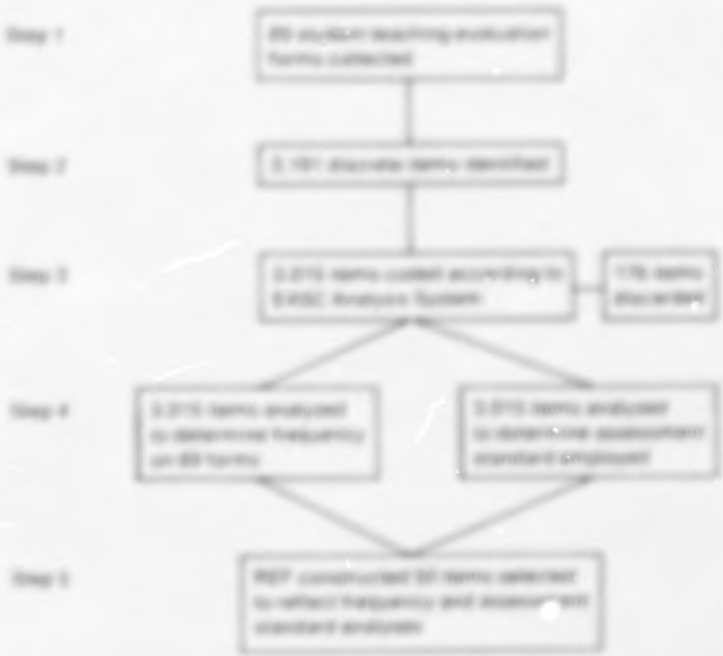
Standard	Code	Explanation/Example	Per Cent
Category	C	The name of an area in which an assessment is to be made: "Questioning."	14%
Definition	D	The name of an area in which an assessment is to be made and mention of the distinguishing characteristics of the area: "Subject matter comprehension: the understanding of concepts through teaching."	1%
Subjective Requirement	SR	A statement that requires a subjective judgment about achievement: "Effective use of blackboard."	73%
Subjective Criteria	SC	A statement that requires a judgment about a level of achievement (scale required): "Acceptance of constructive criticism."	7%
Performance Requirement	PR	A statement that requires rating based on behavioral occurrences: "Accurate reports handed in on time."	6%

Based on the two analyses, 50 actual evaluation items were selected to constitute the REF. All subcategories that had appeared at least once or more than 50 percent of the 89 forms were represented according to their relative frequency. The 50 REF items also reflected the percentages determined by the second analysis. For example, "subjective requirement" was found to have been employed in 73 percent of all evaluations; therefore, 73 percent of the REF items employed a subjective requirement as an assessment standard.

BEST COPY AVAILABLE

Figure 21.1, below, graphically represents the processes involved in the study, as described above, from the collection of the student teaching evaluation forms to the construction of the REF.

Figure 21.1

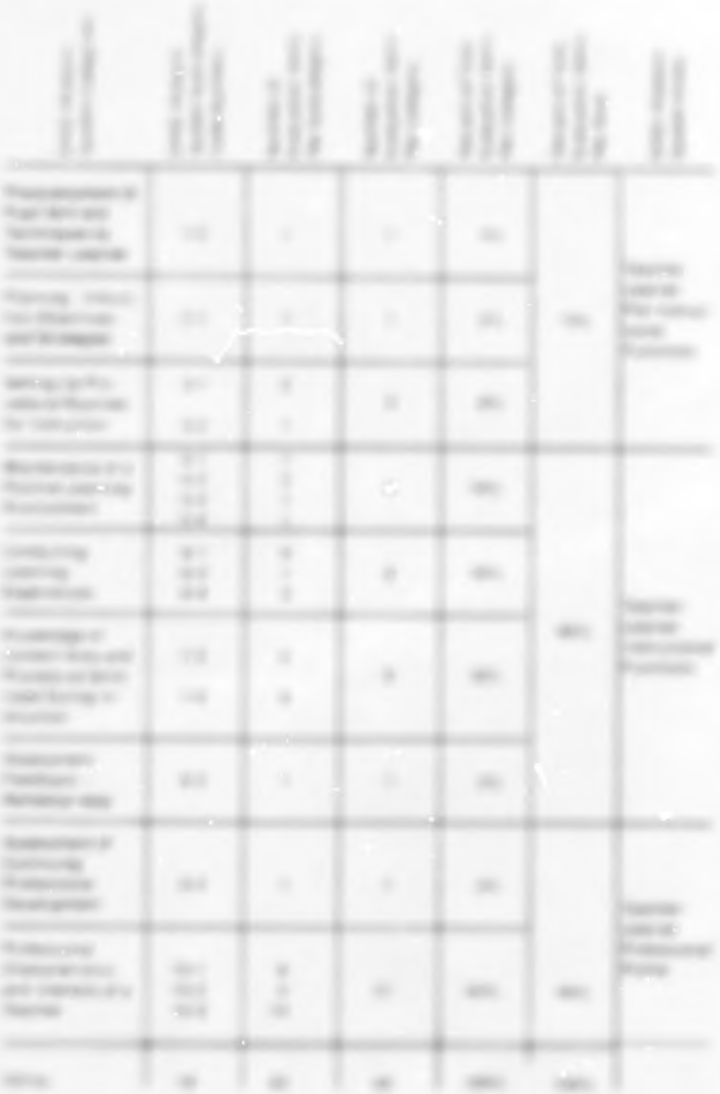

Step 1
89 student teaching evaluation forms collected
Step 2
3,191 discrete items identified
Step 3
3,015 items coded according to EASE Analysis System
176 items discarded
Step 4
3,015 items analyzed to determine frequency on 89 forms
3,015 items analyzed to determine assessment standard employed
Step 5
REF constructed 50 items selected to reflect frequency and assessment standard employed

The following charts highlight the primary findings of the Bunkari study. Chart 21.1—Emphasis on Areas of Assessments in REF Determined by Frequency of Item Occurrence—presents information about the emphasis placed on various portions of the REF. The number of evaluations per sub-category and category and the percent of the whole represented by these figures by category and area are presented. Chart 21.2—Emphasis on 'Subjective Requirement' and 'Subjective Criteria' Assessment Standards in REF Determined by Frequency of Use—presents the number of evaluations in one category that employed either a 'subjective requirement' or 'subjective criteria.' The percent of evaluations in one area that employed either a subjective requirement or subjective criteria has also been calculated.

BEST COPY AVAILABLE

Exhibit 4

EMPHASIS ON AREAS OF ASSESSMENTS AS NOT DETERMINED BY FREQUENCY OF TIME OCCURRENCE



BEST COPY AVAILABLE

2. Teacher lecture and demonstration techniques
3. Questioning and answering skills
4. Instructional materials, visual aides, audio-visual usage skills
5. Supervising pupil skill practice
6. Interactive problem solving, critical thinking, and discussions
7. Relational learning experiences
8. Summary and conclusion: reinforcement and review

7. KNOWLEDGE OF CONTENT AREA AND PROCEDURAL SKILLS USED DURING INSTRUCTION CATEGORY
   1. Accuracy of information
   2. Adequacy in employment of procedural skills
   3. Communication techniques
   4. Utilization and comprehension of principles in content area, understanding basic concepts
   5. Provision of application procedures necessary for pupil problem solving
   6. Understanding or applying educational philosophy
8. ASSESSMENT-FEEDBACK-REMEDIAL HELP CATEGORY
   1. Assessment: pupil product, process, and knowledge for diagnostic purposes after instruction
   2. Individual and group evaluation of pupil work
   3. Feedback and analysis of pupil performance
   4. Remedial help based on analysis of pupil performance
   5. Re-assessment of pupil learning, evaluative instruments, and lesson planning based on feedback and remedial help

TEACHER-LEARNER PROFESSIONAL PROFILE AREA

9. ASSESSMENT OF CONTINUING PROFESSIONAL DEVELOPMENT CATEGORY
   1. Preassessment of performance capacities of teacher learner
   2. Self-analysis and self-evaluation by teacher learner
   3. Openness to critical comments by others
   4. Capacity to identify means of improvement
   5. Pupil learning results: achievement and gain
   6. Pupil reaction to teacher learner
   7. Demonstrating ability to progress as teacher learner
10. PROFESSIONAL CHARACTERISTICS AND INTERESTS OF A TEACHER CATEGORY
    1. Role awareness and ethics
    2. Personal and professional characteristics
    3. Affective capacities enthusiasm, desire, sympathy
    4. Personal involvement in content area activity
    5. Teacher interests: extra-curricular activity, community activity, etc.

BEST COPY AVAILABLE

B

A NEW STYLE CERTIFICATION INSTITUTE: DATA BANK TRAINING INSTITUTE*

Introduction

The purpose of this Data Bank is to collect critical kinds of information at a low cost, which will be of maximum value for providing pertinent information to the New York State Education Department, school districts, and colleges about the performance and effect of teacher education programs on pupil learning in the public schools. Two types of information have been designated as of utmost importance. The first is the final evaluation forms used by supervisors to assess teacher learner's performance in the classroom. The second is an observation by trained observers of the range of these teacher learners' education activities during their training program. Using this information, the Data Bank charts an institutional profile of the training program's means of assessing student teachers and of their actual activities as learners in the classroom. This profile is used to review, evaluate, and formulate instructional programs in ways which increase the decision-making responsibilities and accountability of associated administrators, faculty, and students. The purpose is to increase the range of learning opportunities provided to the teachers-in-training, and simultaneously to provide a widening range of learning opportunities for public school pupils. The Data Bank System, through a continuous feedback process, allows the training staff at the college and public school level to identify gap areas in instruction and to concentrate in the field on filling them. Once the gap is identified, instructional resources are developed which will enable the teacher learner to provide pupil instruction in this neglected area. The formulation of instruction that is adequate in meeting Data Bank standards requires two steps:

(1) That the means of assessment is stated prior to instruction, and is formulated in terms which are observable;

(2) That instruction is provided which fulfills the requirements of these means of assessment. This provides for the educational rights of the teacher learner as adopted by the Advisory Board of the Regents of the State of New York.

Preassessment of Local Institutional Profile

A distinguishing feature of our program is that there is a preassessment profile made of the participating institutions' educational program. This is followed by inservice training based on special needs of the participating institution, in order to provide for immediate positive actions in areas of

*The Data Bank for the state of New York Division of Teacher Education and Certification was designed by the Educational Assessment Systems Corporation, 1432 Amherst Street, Buffalo, New York, 14216.

This system is one of several designs now being considered for adoption by the New York State Education Department.

BEST COPY AVAILABLE

need. When desired, a followup profile is provided to determine areas of gain as a result of the work done. This pattern is the model established by the Data Bank for the certification of programs in the State of New York.

The profile will consist of a 2-day visit to the participating institutions, consisting of:

(1) Sixteen to twenty 30-minute observations made by the EAS staff of the institutions, paraprofessionals, and student teachers.

(2) A collection of information concerning:

a. Student teaching and other teacher learning assessment instruments.

b. Task descriptions of teacher learners' jobs, plus role descriptions, method course outlines, all available instructional modules, and evaluative procedures.

c. Fifteen-minute interviews with 10 student teachers, four cooperating teachers, four method instructors, four supervisors, and administrators to determine a picture of the institutions' governance operations and advisory board, curriculum procedures, program sequence, and management systems for providing feedback on program effectiveness.

The material collected here is necessary in order to provide highly relevant kinds of training experiences which reault in immediate program improvements.

After this preassessment visit and development of the institutions' profile, the necessary resource material is organized as follows for the participating institutions' particular needs.

Assessment of Teacher Learners

Training in assessment techniques for teacher learners will take between 4 and 5 hours. It will consist of a general session on theory concerning the relationship of the "Educational Rights of Students" to the means of assessing their performance. There are two major contents to be mastered. The first concerns the statement of performance criteria for assessment and a related report on major errors made by 89 institutions on their student teacher evaluation forms. The second concerns the necessary range of assessment items if a full training program is to provide an adequate base for pupil and teacher learning. Exercises in both content skills are provided, including a review.

(1) *Theory.* This section will be the introduction of the "Educational Rights of Students." This will be followed by an analysis showing how means of assessment must be related to means of instruction, if training programs are to be effective. Specific requirements essential to relating means of assessment to means of instruction are given so that job responsibilities of the teacher learner are specified and training in them is provided. Without meeting these two sets of requirements, no program can operate effectively.

(2) *Statement of Performance Criteria.* This area will specify mistakes frequently made by educational institutions in the statement of performance

BEST COPY AVAILABLE

assessments for student teachers. These are the "Thou Shalt Not" exercises. This section ends with the analysis of the participating institutions' own evaluation forms.

(3) *Range of Assessments.* This area includes a report on the gap areas in student teacher training programs of 89 institutions in New York State. Also included will be a statement and definition of the 11 Data Bank categories needed for assessment. The Data Bank categories fall into three major groups: (1) prior to instruction, (2) during instruction, and (3) in addition to instruction. There are 55 total items, 41 of which relate to classroom observation. An exercise will then be given in which participants will match statements of assessment with the Data Bank categories. Another exercise will follow, in which the participants will sort their own institutional items for assessment for student teachers under the Data Bank categories. The results from these exercises will be used to identify gap areas in their own institutional student teacher assessment forms.

(4) *Filling Gap Areas.* In the prior two exercises, the gap areas have been identified and the requirements for stating performance criteria clarified. This meeting deals with the formulation of two necessary forms of assessment: convergent types of assessment for which pass/fail types of requirements can be stated, and divergent types of assessment problems which require the formulation of criteria. Success in formulation of both types of means of assessment depend upon identifying specifically the function to be performed by the learner or the elements of the problem to be solved, which distinguish between varying degrees of achievement. This is a critical skill because it determines the requirements of instruction. The final exercise in this session will involve the participant in stating task descriptions for criteria. Included in this exercise is the identification of related job specifications for this task and the development of proficiency and competency tests for determining if the job has been done effectively by the teacher learner. (80 min.)

(5) *Anonymous Evaluation of the Session by the Participants.*

Introduction to the Observation System

Training in the observation system will take between 4 and 5 hours. It consists of a lecture demonstration and video tape exercises of actual classroom situations to show how the observation system is used. The training will also show how to make an analysis of the observation findings. This exercise informs the observer concerning how the observation system provides a basis for necessary decision making regarding the implementation of pupil and teacher learning. The observation system records 41 discrete kinds of classroom activities and skills. They are grouped into nine sections.

(1) *Preassessment.* Determines participants' prior knowledge of observable teacher tasks. Group analysis of the preassessment test will follow for review.

(2) *Goals and Means of Presentation.* Explains the objectives, requirements, and criteria of the observation training session. This will be the means by which the participants will evaluate the success of the workshop.

(3) *Presentation.* Lecture demonstration by experienced staff identi-

BEST COPY AVAILABLE

fying the #1 items on the observation form. A videotape example of each item will be used to supplement the definition.

(4) *Practice.* Consists of two different observation exercises, viewed on video tapes from two different classroom and content areas. Participants will use the observation form for the first time, in order to familiarize themselves with the pattern of using this form. These exercises provide for the recognition of the same item in different content areas. (30 min.)

(5) *Assessment as Analysis.* Consists of simulated field training. The participants will be broken down into two groups. Each participant in the group will observe four different video tapes of actual classroom situations using the observation form as they would in the field. The tapes will be run continuously. (60 min.)

(6) *Evaluation and Feedback.* Using the observation forms done by the participants in field training, a reliability check for each tape will be made. Items will be corrected, and uniformly missed items will be replayed and reviewed. (45 min.)

(7) *Recommendations and Decision Making.* Using one of the participants' corrected observation forms for each of the eight tapes, a composite form will be made on a transparency and the items totaled. The pattern of gaps and strengths will be analyzed. Using the pattern, the participants will then be required to write a recommendation to an advisory board as to what actions should be taken in order to strengthen pupil and teacher learning in gap areas.

One participant will be chosen to act as the person making the recommendations to the board. The other participants will be assigned various roles as board members, such as student teacher, principal, community leader, or director. The board will debate the recommendations and determine the decisions to be made relating to pupil and teacher learning. The following example of an actual memo is the type of decision that can be made. (45 min.)

(8) *Reassessment.* For the final reevaluation the participants will observe themselves going over the entire observation system by viewing video clips of themselves as learners during this program. They will use the observation form to determine the range of instruction that has been provided. (30 min.)

(9) *Anonymous Evaluation of the Session by the Participants.*

Development of Instructional Programs

The institutional profile has now been established in the areas of assessment. Each area of assessment must be matched to an area of instruction. Some missing areas (gap areas) need to be developed. This session will show how to develop and improve existing instructional modules in these areas.

An instructional module must have certain essential components, such as: (1) type of skill or task (knowledge-lecture) to be performed, with the characteristics broken down; (2) demonstration showing alternate ways of performing that task; (3) a one-page analysis of the major elements in the demonstration; (4) an exercise for the participant; (5) video feedback; (6) a retrial; (7) proficiency test with a pass/fail standard establishing that

BEST COPY AVAILABLE

the teacher learner is ready to perform the taught skill in the field; (B) specific requirements or criteria for competency which will be used by supervisors in the field; and (C) an evaluation of the module by the teacher learners to be made immediately after the instructional session.

The cluster of the instructional modules constitutes the curriculum of the training program. From this curriculum, a teacher learner profile is developed.

(1) Exercise. The participants will view capsule modules and rate them according to the nine module components and related criteria. Volunteer participants will do a capsule module on a familiar, everyday subject. These will be rated by the other participants and rant solution for missing areas will be developed with the group.

(2) Practical Exposure. The participants will view two practice video clips of modules in gap areas already tested in the field. (30 min.)

(3) Reenactment. The participants will break into two groups and each will make a minimodule in a priority area. They will learn their module to each other and evaluate it for means of improvement. This represents the prior exercise for developing a full scale module. (75 min.)

(4) Practical Experience. The participants will act as an advisory board in order to designate tasks and necessary work. A framework will be established for maintaining a teacher learner profile including their students' records from the beginning to the completion of their certification program.

An outline of the instructional sequence necessary for achievement can then be determined on this profile with respect to its specific elements as a curriculum. A timetable for the development of needed elements in the instructional program will then be formulated.

At this point, using all the data from the three training sessions, a new instructional profile will be recognized as a new responsibility. (60 min.)

(5) Anonymous Evaluation of the Session by the Participants.

BEST COPY AVAILABLE

22 Implementation of CBTE - Viewed as a Developmental Process*

Dean E. Hall

Innovation adoption in educational institutions is an activity that is often approached with little insight concerning and anticipation of potential consequences. It is an all too common occurrence to observe an air of casualness that is alarming to those especially exposed to change. Often it seems that decision makers decide to adopt one of the recently developed complex innovation bundles (e.g., individually guided education, competency-based teacher education) on impulse, as they would change a technical detail, which is a change that should not always be taken lightly. Adoption of innovations in educational institutions is not a simple phenomenon or singular event. Adoption of educational innovations is a process that generates inertia and, in many cases, antagonistic unexpected outcomes. Innovation adoption can consume much of the energy of the people involved and can cost a great deal in terms of resources dollars, personnel productivity, and not least of all, time.

In this paper, an attempt will be made to share with CBTE program managers the experiences, theory, and research findings of the CBAM project staff of the Research and Development Center for Teacher Education. This staff has been actively involved in studying the CBTE program development and adoption processes in teacher training institutions around the nation. Out of this work and experience a model of "innovation adoption" in educational institutions has been developed. Based on this model, measurement procedures and principles for "adoption agents" are being developed and

*The writing of this manuscript was undertaken under contract with the National Institute of Education. The opinions expressed are those of the author and do not necessarily reflect the position or policy of the National Institute of Education and no endorsement by the National Institute of Education should be inferred.

[illegible]

1. [illegible]
2. [illegible]
3. [illegible]
4. [illegible]
5. [illegible]
6. [illegible]

[illegible]

[illegible]

[illegible]

BEST COPY AVAILABLE

One premise underlying the CBAM is that adoption agents (specialists in the use of the innovation and effective catalysts for facilitating change) work with people in the user system both individually and in groups. As a result, the CBAM at one level focuses on assessing the current state of the individuals within the user system. This assessment then allows the adoption agent to focus his interventions so that they respond to the perceived needs of the individual users and also relate to their levels of use of the innovation at that time. The constructs of the CBAM that make this possible are the two sets of scales:

1. Levels of use of the innovation
2. Stages of concern about the innovation

In addition, a third and more specific set of hypotheses has to do with the relationship of stages of concern to levels of use.

Levels of Use of the Innovation

We contend that there are observable differences in how various individuals approach and use an innovation. Specifically, it is hypothesized that there are identifiable, definable, and measurable levels of use of an innovation that range from lack of knowing that the innovation even exists into an active, sophisticated, and highly effective use of it. It is further hypothesized that improved use of the innovation by most individuals is a developmental process. Normally, individuals do not use an innovation for the first time, or even the second time, as efficiently and effectively as do those who have been involved with the innovation through four or five cycles of use. Advanced levels of use are not attained merely by repeated use of the innovation, however. Experience is essential but not sufficient to insure that a given individual will develop the ability to use an innovation effectively.

An oversimplified but helpful illustration of the level-of-use dimension is the process a college instructor goes through when he adopts a new textbook for a course he has taught many times. At first he will use the new book for reference much more than he did the old text. In preparing class presentations and examinations, he will refer to it much more. His assignments are more likely to be specific chapter assignments, and he probably will follow a straightforward progression through the text. His use of the innovation is apt to be mechanical, uneven in flow, and closely related to the text. As this instructor prepares to teach the course a second time using the new text, however, he is likely to select a different arrangement of assignments. This time, he may assign Chapter 4 first and, perhaps, delete Chapter 7 while substituting another reference he thinks will do a better job. In making these changes, he has progressed beyond a mechanical use of the innovation. He has gained the experience and know-how to be more adaptive in his approach, and he more smoothly integrates the use of the text into the rest of his instructional activities.

For research, operational definitions and scale points for the levels-of-use dimension of the CBAM have been developed. To illustrate, Table 22.1 contains sample behaviors found at each level. Two subscales are hypothesized for the levels-of-use dimension. One describes the *knowledge* level of the

BEST COPY AVAILABLE

user. It hypothesizes that the cognitive level, or amount of information and degree of understanding an individual user has about the innovation, is a developmental progression. Assessment of this set of scale points might take the form of a pencil-paper achievement test. The other set of scale points for level-of-use of the innovation are the *action* scale points. In the CBAM, one hypothesizes that there are observable behavioral differences in how the innovation is actually used and that advancement to the higher levels of use of the innovation is a developmental process. Assessment of the action level of use requires direct observation of the users while they use the innovation.

Table 22.1

LEVELS OF USE AND TYPICAL BEHAVIORS
FOR EACH LEVEL OF USE OF THE INNOVATION

Level of Use	Behavioral Indices of Level
Non-Use	No action is being taken to learn about new ideas in the area of the innovation.
Orientation	The user is seeking out information about the innovation.
Initial Training	The user is preparing to use the innovation.
Mechanical	The user is using the innovation in an awkward, poorly coordinated manner.
Independent	The user is doing a good job with the innovation.
Integrated	The user is sharing with others what he is learning about students from us.ng the innovation.
Renewing	The user is seeking out more effective alternatives to his established use of the innovation.

Stages of Concern About the Innovation

A second dimension has to do with the individual user's needs, motivations, problems, and requests as he is becoming expert in using the innovation. In a way that parallels Fuller's (1969) studies of concerns of teachers, individuals are hypothesized to have concerns that relate to their potential or actual use of an innovation. A set of scale points—Stages of Concern About the Innovation—has been defined for this dimension, and it is hypothesized that this dimension is also a developmental progression. That is, when the individuals first approach using an innovation, their concerns will be different from those they will have after they have used it a while. Still higher stages of concern will be expressed with subsequent cycles of use unless one or more developmental processes become blocked or dormant.[5]

BEST COPY AVAILABLE

As with Fuller's theory of concerns of teachers, the CBAM hypothesizes that early concerns are much more self-oriented than are later concerns. Table 22.2 lists Stages of Concern About the Innovation ranging from Unaware to Renewal with typical expressions of concern.

Table 22.2

STAGES OF CONCERN AND TYPICAL EXPRESSIONS OF CONCERN ABOUT THE INNOVATION

Stage of Concern	Expressions of Concern
Unaware	I don't know anything about it (the innovation).
Awareness	I have heard about the innovation, but I don't know much about it.
Exploration	How much of my time would use of this innovation take?
Early Trial	I seem to be spending all my time in getting material ready for students.
Limited Impact	I can now see how this innovation relates to other things I am doing.
Maximum Benefit	I am concerned about relating the effects of this innovation with what other instructors are doing.
Renewal	I am trying a variation in my use of the innovation that looks like it is going to result in even greater effects.

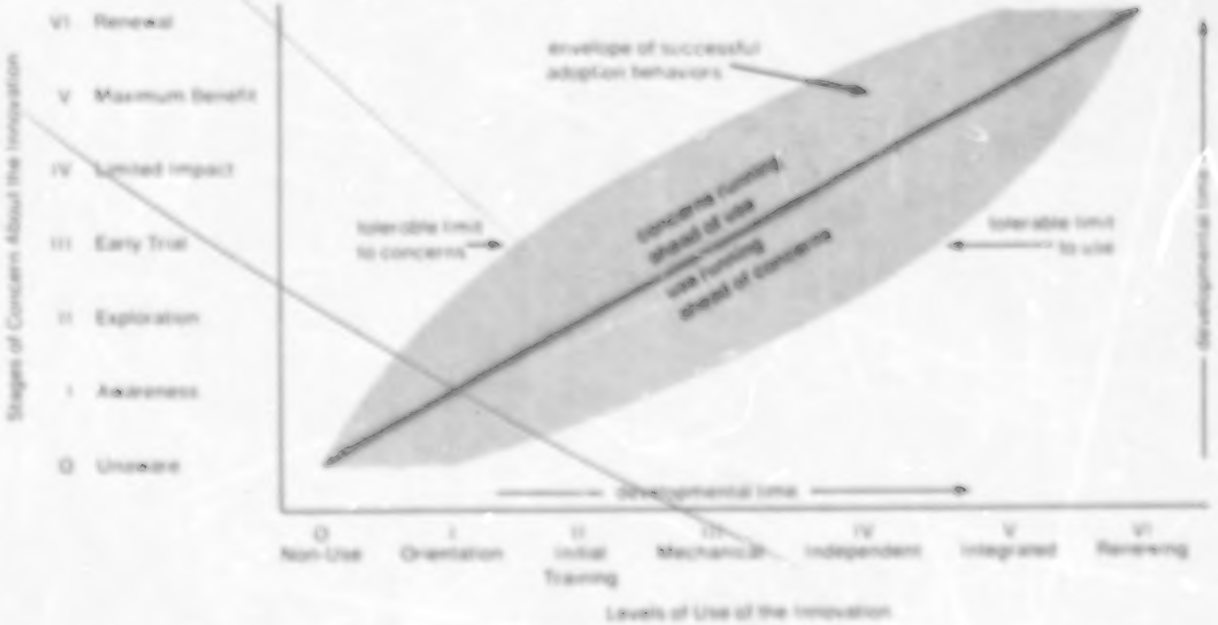
Relationship Between SoC and LoU

It is hypothesized that concerns are related to use and that it is possible for change agents to infer a great deal about use of the innovation from listening to the user's concerns. This relationship is not always a simple one-to-one correspondence, however. Many of us, for example, have known golfers who "talked a good game" but whose actual play was rather far over par. The alternate imbalance in theory is also possible where the individual's concerns are of a very low level and he has serious doubts about his abilities when, in fact, he has the potential of being outstanding. There are also instances of individuals who "perform over their heads."

An illustration of these relationships using an educational innovation could be schools adopting open concept classrooms. Many communities now have school buildings that are open concept and have reputations for having exciting, innovative programs. When one visits some of these schools, however, he finds that bookcases, chalk boards, easels, and seating are arranged in blocks that serve as traditional self-contained classrooms (low

Figure 22.4

RELATIONSHIPS BETWEEN STAGES OF CONCERN AND LEVELS OF USE WITH SUCCESSFUL ADOPTION



BEST COPY AVAILABLE

BEST COPY AVAILABLE

user. In another school, where all the walls are gone, the pupils are roving aimlessly. No territories have been established, there is excessive confusion, and the climate feels tense. This is a school that is probably early in its use of open-concept classrooms and where the teachers have high stages of concern about sharing their leadership and responsibility for curriculum and about remaining nonauthoritarian. But, in spite of these high concerns, their level of use of open-concept classrooms is low. They are confused and uncertain as a result of perhaps attempting too ambitious a beginning.

With the CBAM, it is hypothesized that there is probably a middle range of relationship between concerns and use where successful advancement or growth is possible. But if an individual's stage of concern and level of use move too far out of correspondence, then adoption of the innovation is in jeopardy. Figure 22.4 is a graphic representation of this set of hypotheses, with the area within the envelope representing the hypothesized safe-growth area.

Summary

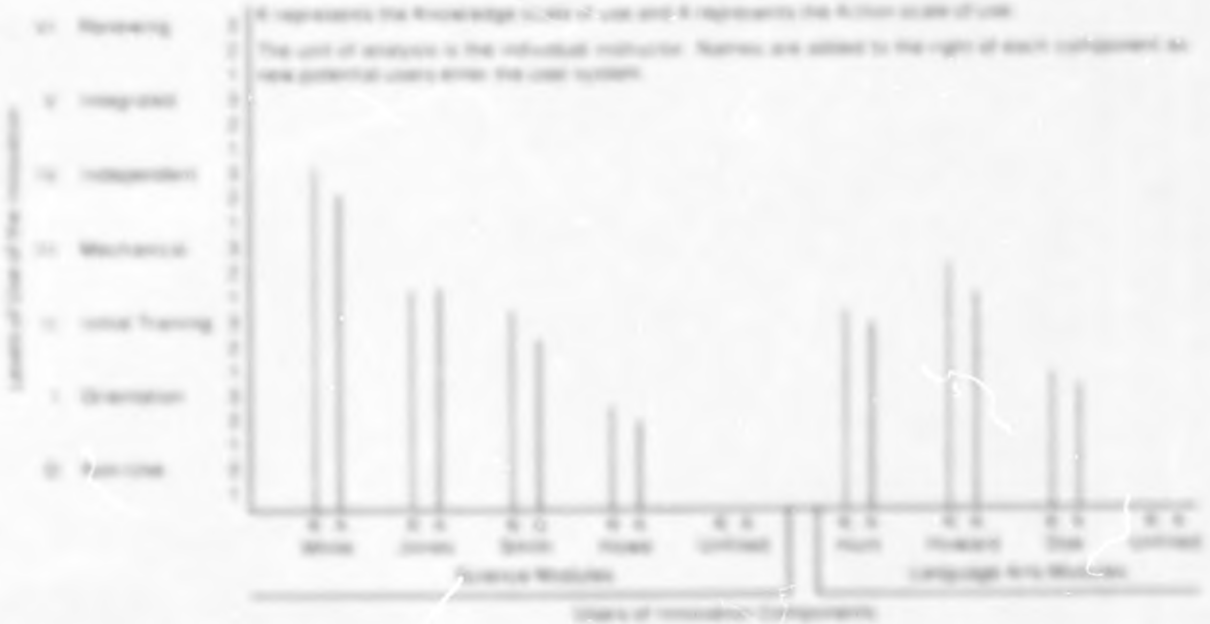
The ultimate criterion in any innovation adoption effort is the extent and quality of use by each user of the innovation within the user system. The Habit-of-use dimension of the CBAM contains a set of operationally defined score points that provides behavioral indicators of the quality of use of an innovation by each individual within the user system. Innovations are adopted by user systems composed of many individuals. It is important to have a record of each individual's level of use. Also, a representation of the proportion of individuals within the user system that are using an innovation needs to be made. A descriptive statement that the average user in a school is at a mechanical level of use is not as useful as is a picture of the present level of use that each individual is demonstrating. An extensity profile can be constructed to accomplish this. All faculty, administrators, and students can be observed and rated with respect to their levels of use of the innovation. When this information is plotted, the resultant graph represents the individual level of use and the extent of use of the innovation within the user system at the time the observations were made. By plotting extensity profiles at regular intervals, a visual record can be maintained of the extent and level of use of the innovation. When extensity profiles for different dates are compared, the rate of advancement of innovation use or its arrest can easily be seen. Figure 22.5 is an example of an extensity profile.

Putting It Together

The Concerns Based Adoption Model in its entirety is represented in Figure 22.6. In operation, there is a collaborative linkage established between a user system that is adopting an innovation and a resource system that has expertise with the innovation and facilitating its adoption. In theory, linkage is accomplished via several communication channels that entail systematic

Figure 33.6

EXTENSITY PROFILE FOR UNIVERSITY X
AFTER TWO YEARS EXPERIENCE IN USING INSTRUCTIONAL MODULES



[illegible] of how each product plan is [illegible] to [illegible] each [illegible] of [illegible] each [illegible] of [illegible] the [illegible]. [illegible] [illegible] [illegible]

[illegible]

[illegible]

[illegible]

AUTHORS